

Систематическая ошибка отбора — статистическое понятие, показывающее, что выводы, сделанные применительно к какой-либо выборке, могут оказаться неточными вследствие неслучайного неправильного отбора объектов в эту выборку.

Типы систематических ошибок

Пространство:

- Выбор первой и последней точки в серии. К примеру, для того, чтобы максимизировать заявленный тренд, можно начать серию с года с необычно низкими показателями и закончить годом с самыми высокими показателями.
- «Своевременное» окончание, то есть тогда, когда результаты укладываются в желаемую теорию.
- Отделение части данных на основе знаний обо всей выборке и затем применение математического аппарата к этой части как к слепой (случайной) выборке (т.н. «предвзвешенная выборка» или «ошибка меткого

– Изучение процесса на интервале (во времени или пространстве), заведомо меньшем, чем требуется для полного представления о явлении.

Данные:

– Вычёркивание неких «плохих» данных в соответствии с придуманными правилами, хотя бы эти правила и шли вразрез с предварительно объявленными правилами для этой выборки.

Участники:

– Пристрастный предварительный отбор участников. К примеру, для доказательства, что курение никак не вредит результатам фитнес-тренировок, можно разместить в местном фитнесцентре объявление для набора добровольцев, но курящих набирать в среди тренеров и опытных участников, а некурящих – среди начинающих или

– Выбрасывание из выборки участников, не дошедших до конца теста (т.н. «*систематическая ошибка выжившего*»).

Во время Второй мировой войны математику Абрахаму Вальду поручили найти варианты бронирования бомбардировщиков: не все они возвращались на базу, а на тех, что возвращались, оставалось множество пробоин от зениток и истребителей на крыльях и хвосте. Значило ли это, что в этих местах нужно было больше брони? Вальд ответил: нет, эти места достаточно защищены. Самолёт, которому попали в кабину или топливный бак, выходит из строя и не возвращается. Поэтому укреплять надо те места, которые наиболее «чистые». «Слухи об уме и доброте дельфинов основаны на рассказах пловцов, которых они толкали к берегу, а не в открытое море». Из старых зданий остаются только самые красивые и прочные — остальные подвергаются сносу, и на их месте строят новые.

Классическое искусство — это то, что проверено временем, но плохими работами того времени большинство людей не

Когортные исследования

Данные исследования обладают большой доказательной силой и широко используются исследователями США и стран Западной Европы. Эти исследования **основаны на наблюдении и не предполагают активного вмешательства** исследователей в естественное течение изучаемого процесса; они также всегда являются «**продольными**», и почти всегда – **проспективными** (бывают и ретроспективные когортные исследования).

Основная цель когортных исследований – установить, **влияет ли определенное воздействие (или несколько) на последующее развитие определенного исхода** (например, клинически манифестного заболевания).

Кроме того, исследования подобного типа предпринимаются для точного описания типичного течения и характерных осложнений различных заболеваний, частоты встречаемости каких-либо событий, интересующих ученых, изучения побочных эффектов и взаимодействия лекарственных препаратов, выявления факторов риска различных заболеваний и состояний и оценки степени их вклада в развитие изучаемого исхода и т.п. Это **лучший вид клинических исследований для тех случаев, когда эксперимент невозможен.**

Принцип когортных исследований – в продолжительном наблюдении за одной (реже – несколькими) группами исследуемых лиц («когортами»). Срок наблюдения варьирует от нескольких месяцев до десятков лет и оговаривается заранее (хотя в дальнейшем может быть увеличен в соответствии с результатами промежуточных анализов полученных данных). Размер когорты обычно стараются сделать как можно большим; в идеале когорта совпадает по численности с исследуемой генеральной совокупностью. В случае, если исследуется широко распространенное заболевание или состояние, представляющее большую социальную значимость, и все случаи таких заболеваний тщательно регистрируются на государственном уровне, описанный выше идеальный вариант достижим; так, ученые ряда стран исследуют национальные когорты ВИЧ-инфицированных или больных атеросклерозом. При согласованном объединении когорт нескольких стран образуются огромные интернациональные когорты, так поступает при организации международных

Когортное исследование отчасти похоже на исследование вида «случай-контроль»:

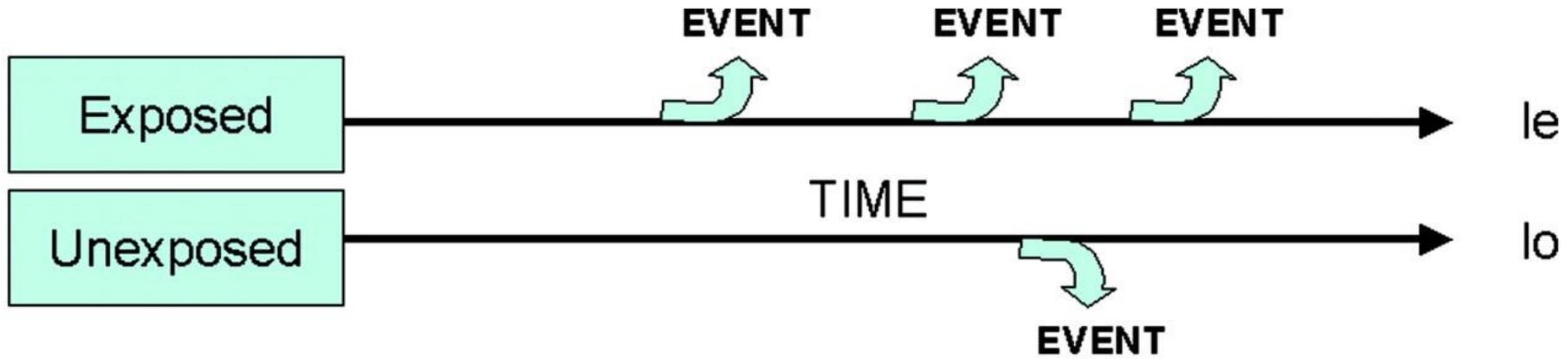
– когорта формируется из лиц, имеющих риск развития какого-либо состояния, интересующего исследователей (например, ВИЧ-инфекции – у взрослых социально и сексуально активных лиц, ОНМК – у пациентов с артериальной гипертензией, обострения/хронизации – у пациентов с вирусными гепатитами, побочных эффектов АРТ – у получающих АРТ и т.п.);

– в дальнейшем в ходе наблюдения регистрируются случаи развития искомого состояния (т.н. «исходы»). По завершении исследования когорта разделяется на группу, где *действовал интересующий исследователей фактор* («опытная группа»), и группу, где *указанный фактор не действовал* («контрольная группа»), после чего в указанных группах сравниваются

– если в ходе когортного исследования стоит цель *просто описать частоту развития некоего явления в определённой группе лиц*, подвергающихся воздействию какого-либо фактора (например, рака лёгких либо ХОБЛ у курильщиков, ВИЧ-инфекции – у парентеральных наркоманов и т.п.), то когорта формируется из лиц, находящихся под воздействием соответствующего фактора (например, из парентеральных наркоманов), а контрольная группа не нужна;

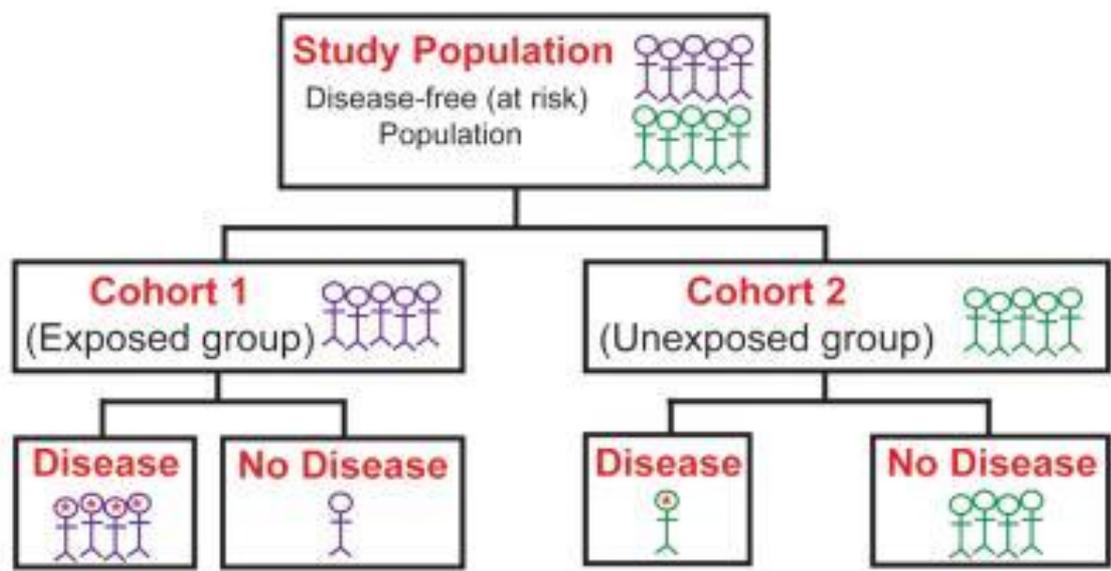
– всё вышеперечисленное должно учитываться в критериях включения/исключения в исследование;

– перед проведением когортного исследования желательно выполнение «пилотного» исследования с менее доказательным типом дизайна для получения начальной информации о структуре изучаемого явления – это необходимо для процедуры

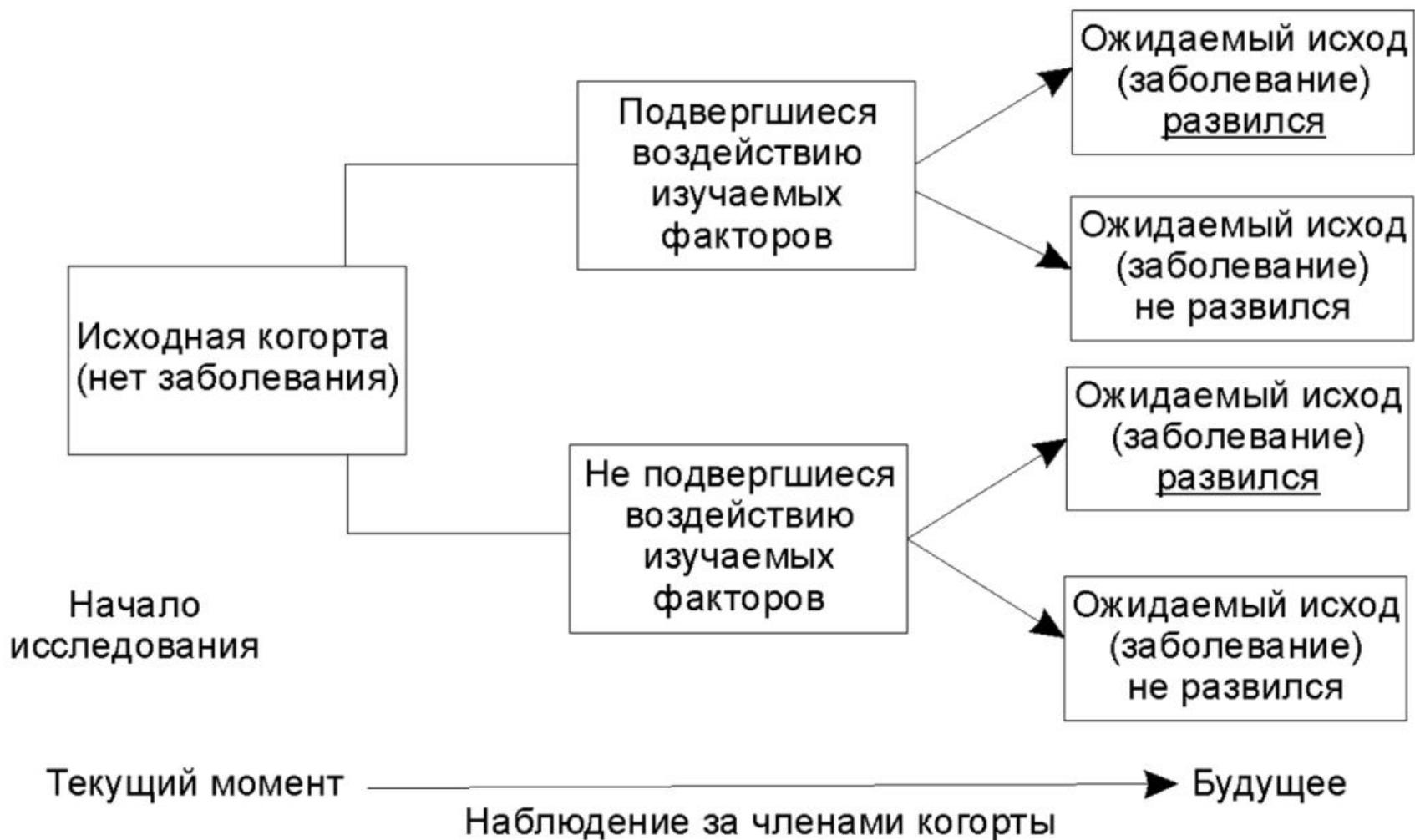


Population At Risk

COHORT STUDY



1. Identify exposed and unexposed cohort groups.
 - Unexposed
 - Exposed
 - Diseased status
- 2a. PROSPECTIVE STUDY
-During follow-up period, identify diseased subjects (incident cases).
- 2b. RETROSPECTIVE STUDY
-Identify diseased subjects by interview or written records.
3. Analyze differences (i.e. incidence or relative risk) among those exposed (cohort 1) and those unexposed (cohort 2).



Категории критериев включения/исключения (исходя из контекста исследования)

Коморбидность

Сопутствующие заболевания (в настоящее время или ранее), не изучаемые в исследовании. Могут быть критериями исключения.

Лечение

Медикаментозная терапия, проводимая ранее или в настоящее время, хирургические вмешательства, участие в другом исследовании.

Форма или степень тяжести заболевания

Клиническая форма и степень тяжести изучаемого заболевания

Критерии, связанные с беременностью

Беременность, лактация, методы контрацепции

Персональные критерии

Диагностические процедуры

Касается выполнения обследований (не их результата!) диагностических процедур, необходимых для включения в исследование (например, измерение кровяного давления, любые специфические лабораторные тесты и т.д.)

Прочее

Например, такие специфические индивидуальные критерии, как исключение при беременности *партнёра*, либо включение в зависимости от места жительства, наличия информированного согласия, грамотности/способности изъясняться на государственном языке и т.д. Также могут использоваться в качестве дополнительных критериев исключения (например, ожидаемая низкая приверженность рекомендациям и т.п.)

Примером является **EuroSIDA**, согласованно проводимая на 6 когортах ВИЧ-инфицированных, проживающих во многих странах Евросоюза, включая Польшу и страны Балтии, а также в России, Белоруссии и на Украине; общий размер объединенной когорты – более 11.200 больных. Другим примером мультицентрового когортного исследования является **D:A:D** (сбор данных о побочных эффектах АРТ). Объединенная когорта данного исследования включает 11 национальных когорт ВИЧ-инфицированных лиц, проживающих в 21 европейской стране, а также в США и Австралии и наблюдающихся в 188 клиниках; суммарный размер данной когорты составляет 23.468 больных. Если в когорту удалось (или планируется) включить всех зарегистрированных больных с изучаемой патологией, она называется **регистром**. Регистр может быть региональный, национальный и (для редких заболеваний) – международный; регистр больных с острым коронарным синдромом Global Registry of Acute Coronary Events (GRACE), его приблизительным аналогом в России является регистр ОКС РЕКОРД.

Чем больше размер когорты, тем точнее получаемые данные и тем ближе они приближаются к таковым в генеральной совокупности; при этом пропорционально возрастает стоимость исследования и сложность его организации. Большие мультицентровые когортные исследования по силам только большим интернациональным научным коллективам при поддержке влиятельных спонсоров, как правило – фармацевтических компаний.

Когортные исследования бывают фиксированные (после достижения определенного размера когорты включение в исследование новых пациентов прекращается) и динамические (возможно пополнение когорты новыми участниками в ходе наблюдения).

Лица, включенные в когорту, находятся под наблюдением до проявления у них изучаемого исхода (outcome, endpoint).

Исходом может быть, например, развитие какого-либо заболевания либо его осложнения, выздоровление или смерть больного, неудача лечения, появление побочного эффекта терапии; такие исходы называют **истинными**, поскольку они отражают реальные клинически значимые события в жизни исследуемого лица. Кроме того, бывают **суррогатные** исходы – обычно достижение определенного уровня каким-либо показателем, полученным при лабораторных либо инструментальных исследованиях (например, уровня CD4+ лимфоцитов, вирус-нагрузки плазмы крови, холестерина и т.п.). Суррогатные исходы косвенно указывают на развитие тех или иных истинных исходов (например, неудачи лечения или обострения заболевания), но сами таковыми не являются. Два или более независимых события, последовательного или одновременного наступления которых ждет исследователь, могут

объединяться в т.н. **композитный исход** (composite) (например

При анализе результатов когортных исследований наиболее часто используется *метод анализа времени до наступления исхода* (также называется «анализ вероятности наступления изучаемого исхода в определенный период времени» или «*анализ дожития*»).

Корректные результаты могут быть получены только при анализе синхронизированной когорты, т.е. такой когорты, каждый из членов которой был включен в исследование в строго определенный и одинаковый для всех момент развития своего заболевания / состояния (наступление определенной стадии заболевания, первичное проявление симптомов заболевания или его обострение, госпитализация, начало лечения, оперативное вмешательство и т.п. Данное условия требует тщательной формулировки **критериев включения участников в исследование**; в дальнейшем эти критерии должны неукоснительно соблюдаться.

Пациенты, включенные в исследование в разное время, могут быть несопоставимы по ряду признаков. Например, в случае хронических инфекционных заболеваний (вирусные гепатиты В, С, ВИЧ-инфекция и т.п.) методы диагностики и лечения достаточно быстро улучшаются с течением времени; если сравнить больных, включенных в исследование 15 лет назад и в настоящий момент, то окажется, что у современных больных заболевание было выявлено на более ранней стадии, а прогноз существенно лучше, чем 15 лет назад, из-за радикального улучшения терапевтических подходов. Систематическая ошибка данного рода, связанная с улучшением методов диагностики и лечения по ходу продолжительного исследования, называется «ошибка временного сдвига» (lead-time bias); примером такой ошибки является исследование выживаемости ВИЧ-инфицированных больных, заболевание у которых было выявлено до и после 1985 г., из которого ожидаемо следует, что у

Аналогично, при исследовании выживаемости ВИЧ-инфицированных лиц на фоне ВААРТ, начатой при различных уровнях CD4+ лимфоцитов (50 клеток/мм³ и 350 клеток/мм³, соответственно) оказалось, что **при начале ВААРТ на фоне выраженной иммуносупрессии (50 клеток/мм³) выживаемость больных существенно ниже по сравнению с исходно высоким уровнем CD4+ лимфоцитов.**

При этом исследователи не учли, что процесс снижения уровня T_H-лимфоцитов с 350 до 50 клеток/мкл занимает несколько (2-7) лет, которые стоило бы приплюсовать к выживаемости лиц с глубоким иммунодефицитом. Данный пример также подчеркивает необходимость синхронизации формируемой когорты по ключевым показателям, способным оказать критическое влияние на ожидаемый результат исследования.

Другой систематической ошибкой является т.н. **ошибка выжившего** (**survivorship bias**). Данная ошибка возникает в случае, если сравнивается выживаемость больных, получавших или не получавших определенную терапию, причем анализируемый метод лечения стал доступен для использования на определенном этапе уже проводимого когортного исследования. В этом случае для того, чтобы подвергнуться определенному лечению, *больные должны были дожить до его введения в клиническую практику*, а все больные, умершие до этого момента, не имели шанса получить указанную терапию. Соответственно, оказывается, что лечение получили больные, у которых прогноз заболевания исходно был наиболее благоприятным, и в результате эффективность оцениваемой терапии может быть существенно завышена.

В целом, для обеспечения корректности анализа выживаемости необходимо соблюдение следующих условий:

- 1) выборка лиц, вошедших в когорту, случайна и репрезентативна;
- 2) наблюдения независимы;
- 3) в период исследования не происходило изменений в методах диагностики, лечения и процедурах наблюдения (т. е. критерии включения не изменялись);
- 4) в период исследования для всех больных, входящих в когорту, вероятность наступления изучаемого исхода не изменялась;
- 5) случаи смерти, выбывания из исследования и включения новых больных в когорту происходили более или менее равномерно на всем периоде наблюдения.

Для оценки влияния каких-либо факторов на развитие изучаемого исхода определяют **следующие численные показатели**:

1) **Риск наступления исхода к моменту X** (risk of event by time t) – некоторое число в интервале между 0 и 1.

$$\text{Риск} = \frac{\text{Число лиц с изучаемым исходом к моменту } X}{\text{Общее количество лиц под наблюдением}}$$

2) **Шанс наступления исхода к моменту X** (odds of event by time t) – некоторое число между 0 и бесконечностью, приблизительно-но равен риску, если частота исследуемого исхода невелика.

$$\text{Шанс} = \frac{\text{Число лиц с изучаемым исходом к моменту } X}{\text{Число лиц без изучаемого исхода к моменту } X}$$

3) **Частота (наступления) исхода** – incidence rate (также просто incidence или просто rate)

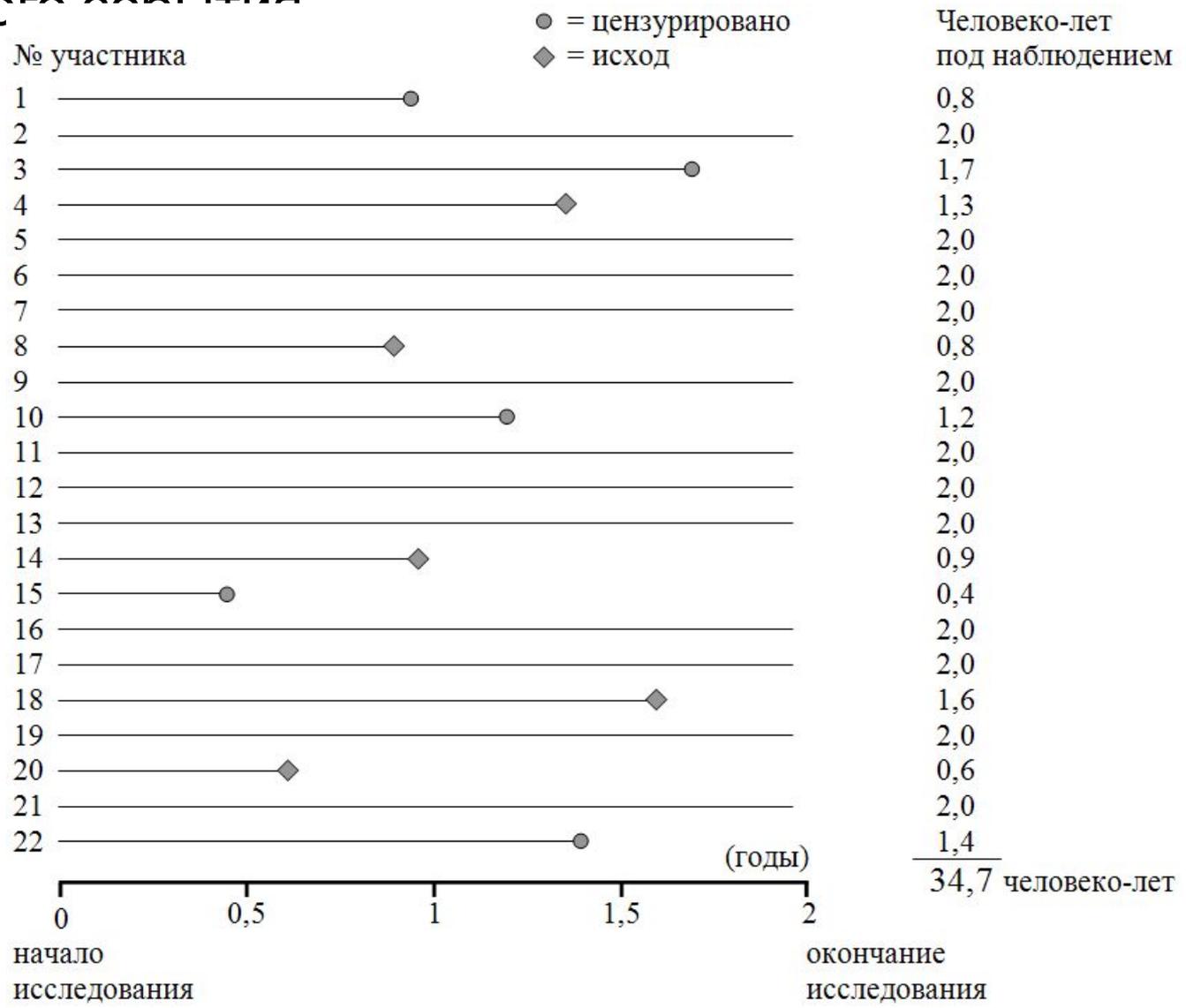
$$\text{Частота исхода} = \frac{\text{Число лиц с изучаемым исходом}}{\text{Сумма человеко – лет под наблюдением}}$$

Определение **частоты исхода** предпочтительнее, чем оценка *риска* или *шанса*, поскольку при этом учитывается тот факт, что не все члены когорты находятся под наблюдением в течение одинакового временного интервала.

Если какой-либо из наблюдаемых индивидуумов вышел из исследования до его завершения по любым причинам (кроме наступления ожидаемого исхода), считается, что время наблюдения в данном случае **цензурировано** моментом последнего планового осмотра данного лица. Такое наблюдение называется неполным, незавершенным или цензурированным.

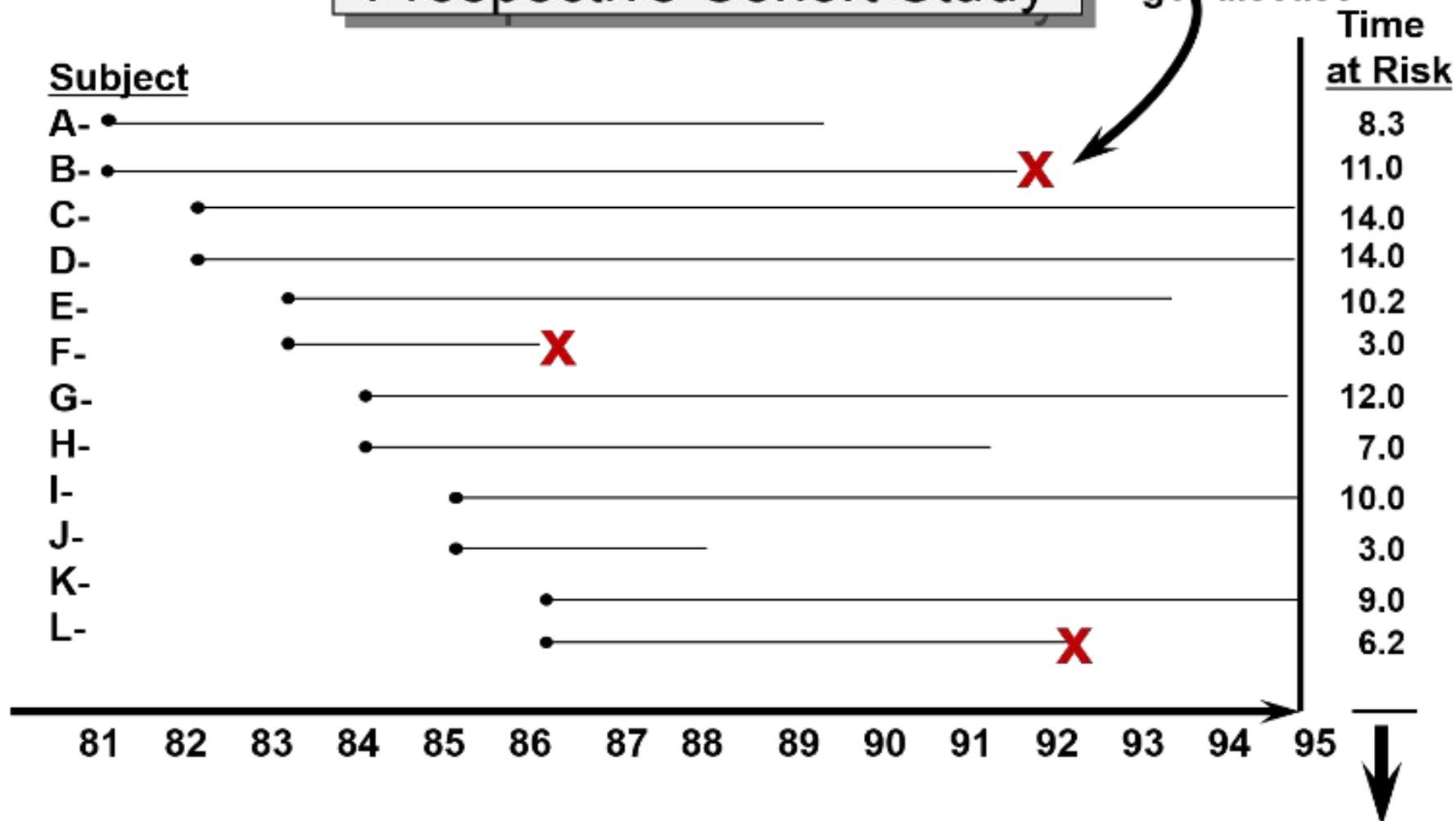
К неполным относятся также те наблюдения, в которых изучаемый исход не наступил на момент окончания исследования. Полным, завершенным или нецензурированным считается такое наблюдение, в котором изучаемый исход наступил до окончания исследования, и при этом точно известен интервал времени между включением больного в исследование и наступлением исхода.

Сумма человеко-лет под наблюдением (person-years at risk) – это сумма времени, проведенного всеми участниками исследования под наблюдением в ожидании наступления изучаемого события



Incidence Rate in A Prospective Cohort Study

X = when they got disease



Incidence (IR) = $3/107.7 = .028/ \text{person-yr}$
 $= 28/1000 \text{ p-yrs}$

Total time at risk = 107.7 person-yrs

Частота исхода – величина постоянная для всего периода исследования, т.к. вычисляется на основании суммарных данных за весь период (так, для ситуации, изображенной на рисунке, *частота исхода* составляет $5 \div 34,7 = 0,14$ на 1 человеко-год).

В то же время ясно, что вероятность появления ожидаемого исхода на разных временных отрезках неодинакова и с течением времени может закономерно изменяться (например, вероятность умереть возрастает пропорционально возрасту, начиная приблизительно с 40 лет). Вероятность развития интересующего нас исхода в данный момент времени характеризуется показателем, называемым «**вероятность опасности исхода**» (**event hazard rate**). Она вычисляется как отношение числа лиц с ожидаемым исходом, развившимся за интересующий нас отрезок времени (определенный день, неделю, месяц, год) к количеству человеко-лет (месяцев, недель, дней), проведенных под наблюдением за это время всеми

Собственно сравнение частоты развития исхода в различных группах внутри когорты, выделенных по наличию или отсутствию воздействия факторов, влияние которых на исход изучается в данном исследовании, производится при помощи вычисления относительных характеристик – **отношения рисков** (risk ratio), **отношения шансов** (odds ratio) и **отношения частот** (rate ratio). Для всех указанных величин рассчитывается 95% **доверительный интервал** (95% CI).

$$\text{Отношение рисков} = \frac{\text{риск в группе 1}}{\text{риск в группе 2}}$$

$$\text{Отношение частот} = \frac{\text{частота в группе 1}}{\text{частота в группе 2}}$$

$$\text{Отношение шансов} = \frac{\text{шанс в группе 1}}{\text{шанс в группе 2}}$$

Результатам когортных исследований о влиянии каких-либо факторов на частоту появления определенного события можно доверять, если:

1) выявленное влияние изучаемого фактора на регистрируемый исход велико (относительный риск, отношение частот, отношение шансов **>2 или <0,5**);

2) аналогичные результаты получены более чем в двух независимых исследованиях;

3) систематические ошибки и явление «смешивания эффектов» при наборе когорт, наблюдении за участниками исследования и анализе результатов отсутствуют, либо систематические ошибки имеются, но в разных исследованиях они имеют различную направленность (т.е. неодинаковы).

Оценка времени до наступления ожидаемого исхода

также называется «**анализ дожития**» (**методика Каплана-Мейера**).

Промежуток времени, лет, прошедший с момента начала исследования	Число лиц под наблюдением без исхода (в данный временной интервал)	Количество <u>развившихся</u> исходов (в данный временной интервал)	Количество неполных наблюдений («censored») в данный временной интервал	Вероятность развития исхода в данный момент (интервал) времени	Вероятность отсутствия исхода в данный момент (интервал) времени	Вероятность отсутствия исхода за все время с начала исследования, вплоть до данного момента
0,1	22	0	0	0	1,00	1,0
0,2	22	0	0	0	1,00	$1,0 \times 1,0 = 1,0$
0,3	22	0	0	0	1,00	$1,0 \times 1,0 = 1,0$
0,4	22	0	1	0	1,00	$1,0 \times 1,0 = 1,0$
0,5	21	0	0	0	1,00	$1,0 \times 1,0 = 1,0$
0,6	21	1	0	$1 \div 21 = 0,048$	0,952	$1,0 \times 0,952 = 0,952$
0,7	20	0	0		1,00	$1,0 \times 0,952 = 0,952$
0,8	20	1	1	$1 \div 20 = 0,050$	0,950	$0,952 \times 0,950 = 0,904$
0,9	18	1	0	$1 \div 18 = 0,056$	0,944	$0,904 \times 0,944 = 0,853$
1,0	17	0	0		1,00	$1,0 \times 0,853 = 0,853$
1,1	17	0	0		1,00	$1,0 \times 0,853 = 0,853$
1,2	17	0	1		1,00	$1,0 \times 0,853 = 0,853$
1,3	16	1	0	$1 \div 16 = 0,063$	0,937	$0,853 \times 0,937 = 0,799$
1,4	15	0	1		1,00	$1,0 \times 0,799 = 0,799$
1,5	14	0	0		1,00	$1,0 \times 0,799 = 0,799$
1,6	14	1	0	$1 \div 14 = 0,071$	0,929	$0,799 \times 0,929 = 0,742$
1,7	13	0	1		1,00	$1,0 \times 0,742 = 0,742$
1,8	12	0	0		1,00	$1,0 \times 0,742 = 0,742$
1,9	12	0	0		1,00	$1,0 \times 0,742 = 0,742$
2,0	12	0	12		1,00	$1,0 \times 0,742 = 0,742$

Из таблицы видно, что *число лиц под наблюдением без исхода* (см. колонку №2) уменьшается всякий раз после регистрации очередного исхода или выхода члена когорты из исследования (регистрации неполного, или цензурированного, наблюдения) на суммарное число лиц, покинувших исследование за рассматриваемый временной интервал (исходы + неполные наблюдения), причем указанное уменьшение регистрируется начиная со следующего за текущим временного интервала.

Размер временного интервала выбирается произвольно, исходя из цели исследования. Обычно его продолжительность устанавливается равной промежутку времени между плановыми обследованиями членов когорты, причем собственно *момент обследования разграничивает соседние интервалы.* Именно поэтому лица, выбывшие из исследования по разным причинам, учитываются начиная с отрезка времени, *следующего за*

Все наблюдения, в которых ожидаемый исход не был зафиксирован к моменту окончания исследования, считаются цензурированными, что и отражено в последней ячейке колонки №4 («количество неполных наблюдений...»). Если участник исследования по каким-либо причинам (кроме регистрации ожидаемого исхода) покинул когорту до окончания исследования, такое наблюдение считается *цензурированным по времени последнего планового осмотра* указанного лица (т.е. последним эпизодом наблюдения за данным членом когорты считается момент последнего планового осмотра, во время которого не было зафиксировано ожидаемого исхода).

После расчета вероятности отсутствия исхода вплоть до момента завершения исследования производится построение **графика Каплана-Мейера**, который отражает вероятность отсутствия изучаемого исхода у членов когорты с момента начала исследования до момента его окончания.

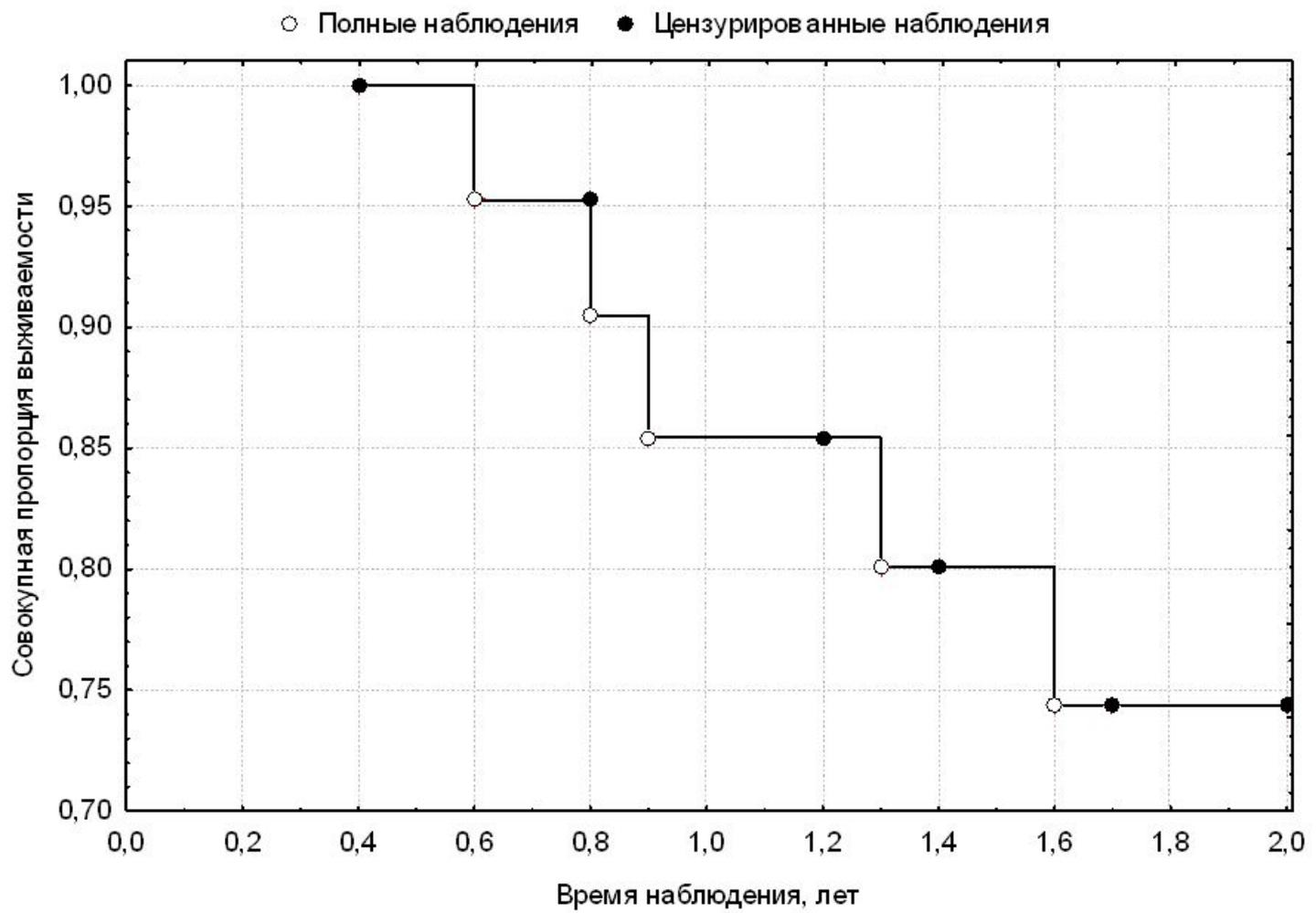


График Каплана-Мейера представляет собой серию «ступенек», где снижение *кумулятивной вероятности отсутствия исхода* происходит скачкообразно после регистрации очередного случая (случаев) ожидаемого исхода; регистрация цензурированных наблюдений не приводит к снижению вероятности отсутствия исхода, вследствие чего «ступенька» на графике не образуется, но сам факт регистрации неполного наблюдения отмечается на графике специальным значком.

Данная особенность построения графиков Каплана-Мейера напрямую вытекает из *принципа организации когортных исследований*, который подразумевает не непрерывный, а периодический учет изменений в состоянии членов когорты во время регулярных контрольных осмотров, вследствие чего изменение вероятности наличия либо отсутствия искомого состояния происходит не плавно, а скачками.

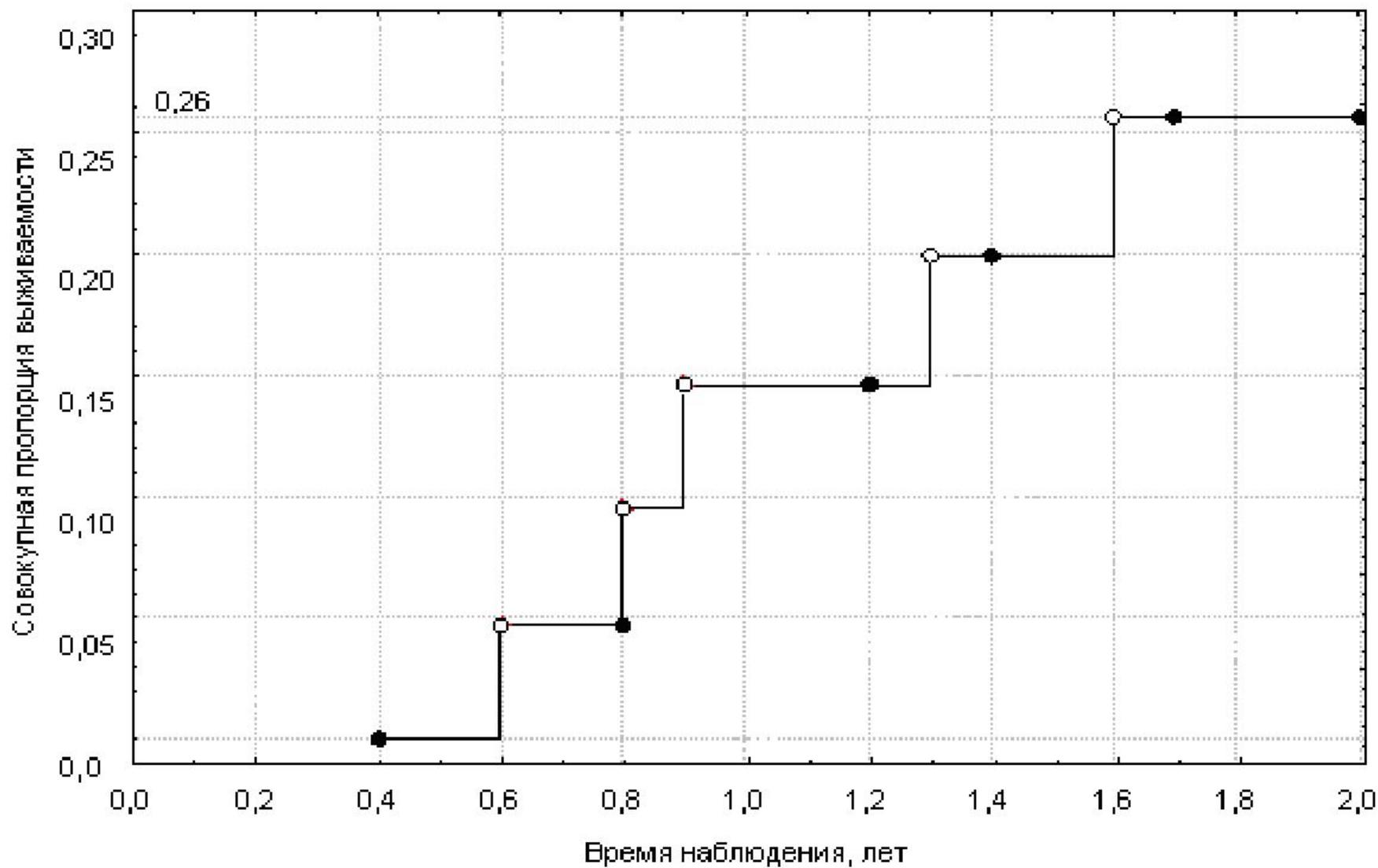
Изображенная на графике Каплана-Мейера кривая называется ***кривой дожития***.

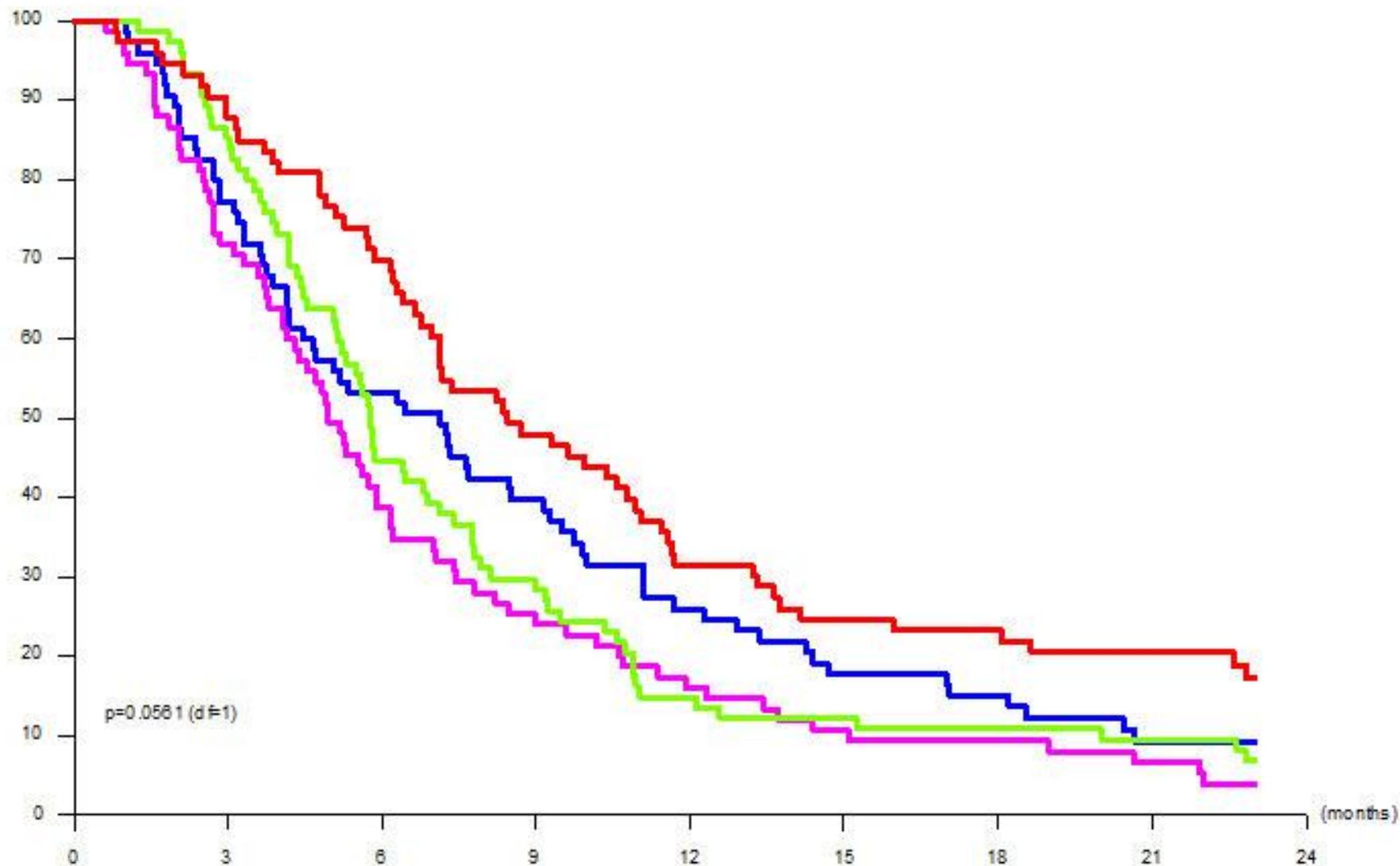
Две кривые дожития, полученные при анализе независимых групп, можно сравнивать, используя **лог-ранговый тест** (log rank test); если в результате вероятность нулевой гипотезы (p) оказывается равна или менее 0,05, то кривые дожития отличаются друг от друга с вероятностью 95% и более.

Кривую дожития можно представить и в альтернативном виде – как кривую вероятности появления ожидаемого исхода с течением времени.

Функция выживаемости

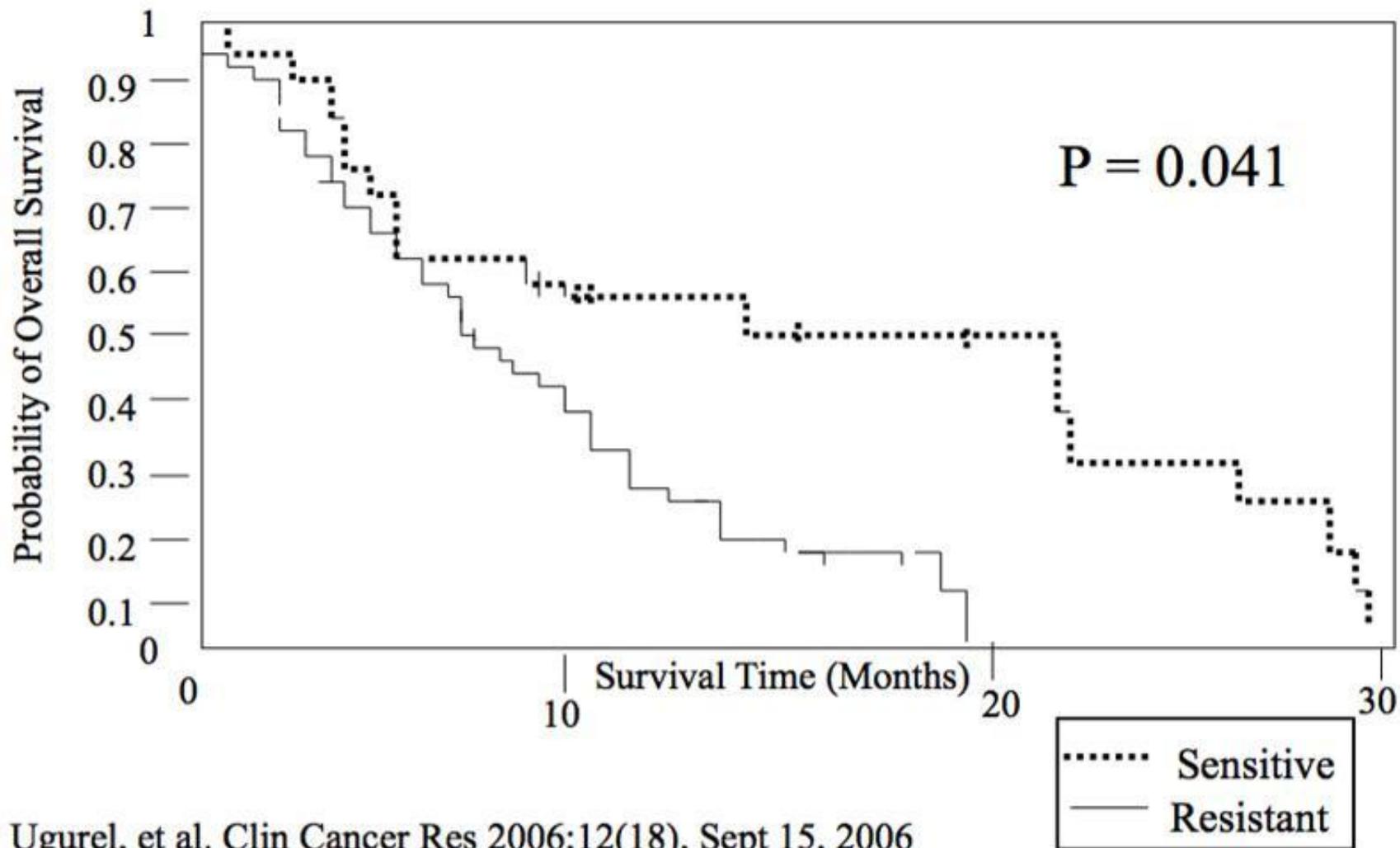
○ Полные наблюдения ● Цензурированные наблюдения





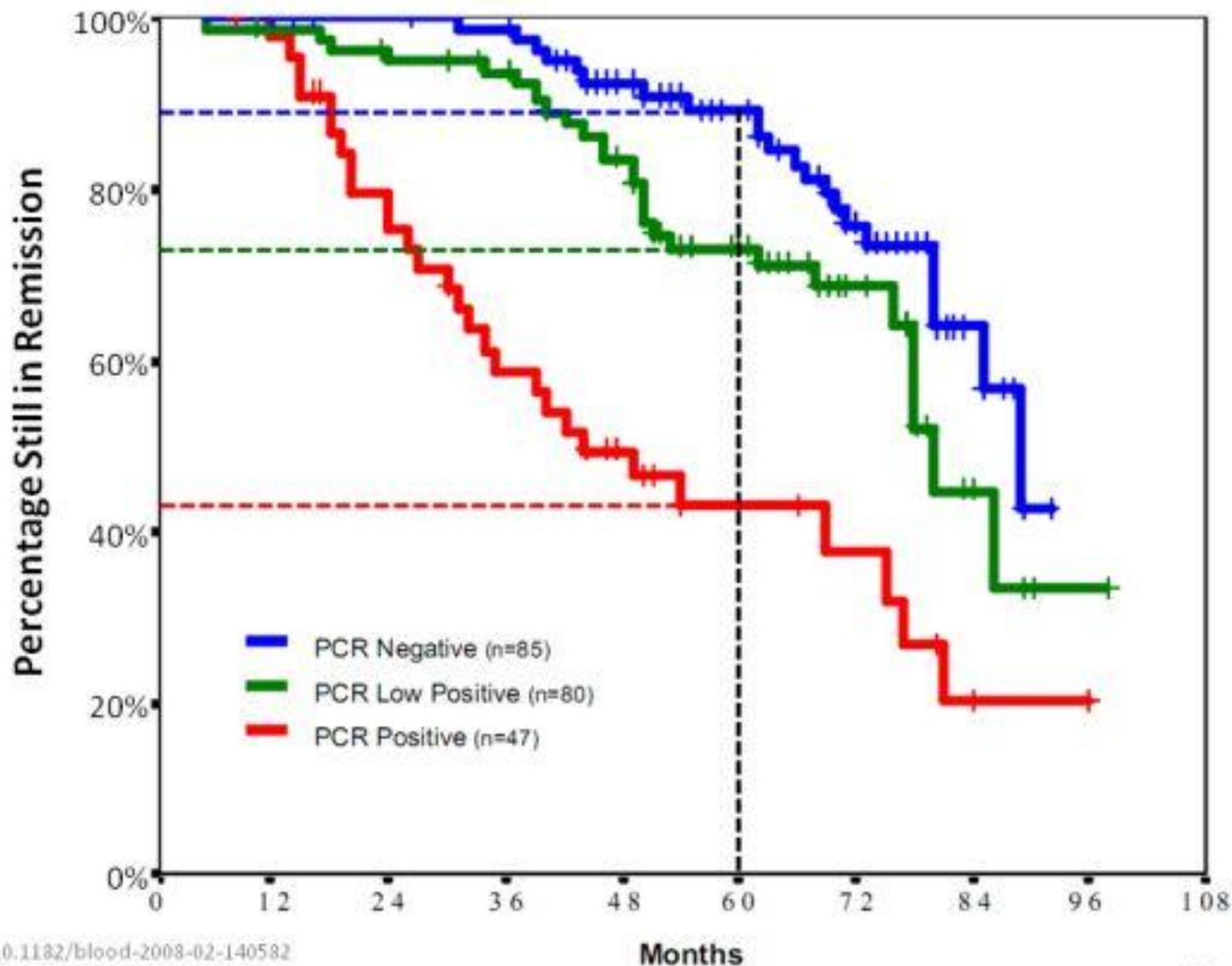
O	N	Number of patients at risk :								Time since first histolo
69	75	58	40	29	19	13	11	6	6	<32.7 wks
73	75	54	29	18	12	8	7	5	5	>=32.7 - <44.1 wks
73	75	64	33	21	11	9	8	7	7	>=44.1 - <65 wks
67	75	64	51	35	23	18	17	13	13	>=65 wks

Melanoma

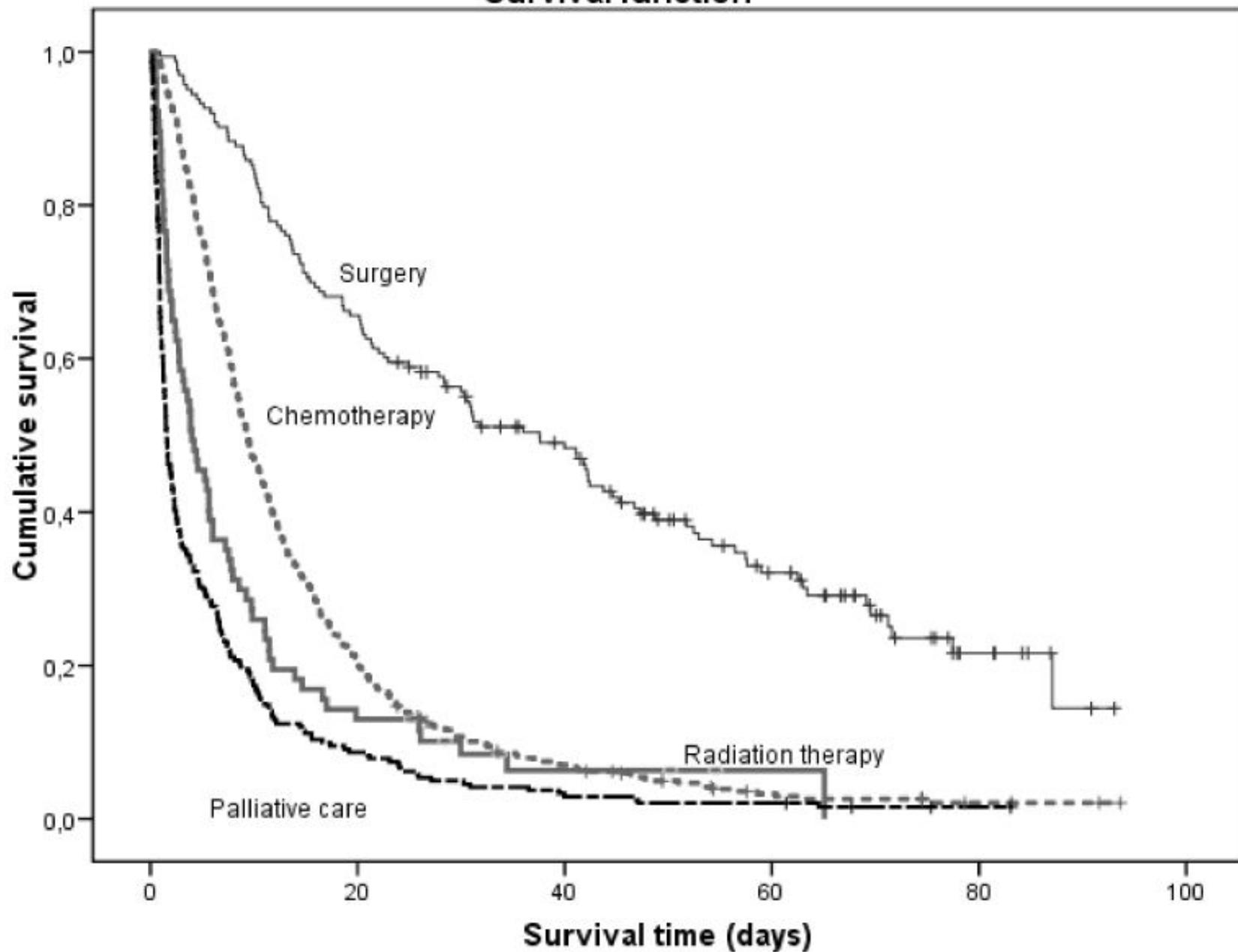


Ugurel, et al. Clin Cancer Res 2006;12(18), Sept 15, 2006

FCR: Duration of Remission



Survival function



Следует обратить внимание на то, что вероятность развития исхода к моменту окончания исследования составляет примерно **26% (1 – 0,742)**, при этом риск наступления исхода к тому же моменту, рассчитанный по приведенной ранее формуле и имеющий тот же смысл, равен **$5 \div 22 = 22,7\%$** .

Величина риска получилась заниженной, поскольку формула для его вычисления не предусматривает постепенное уменьшение размера когорты в ходе исследования (ввиду наступления исхода у части наблюдаемых лиц, а также вследствие выхода больных из состава когорты по другим причинам). Таким образом, **применение методики Каплана-Мейера дает более точные результаты**, чем просто расчет риска наступления события к интересующему моменту, ввиду чего *данная методика является предпочтительной для обработки результатов когортных исследований.*

Помимо построения графика Каплана-Мейера, **анализ дожития подразумевает определение следующих показателей:**

- 1) **дожитие** (доля участников исследования, у которых искомый исход не наступил) для интересующего срока наблюдения;
- 2) **стандартная ошибка кумулятивного дожития** – рассчитывается при помощи соответствующей функции программы для статистического анализа (standard error of cumulative survival), в нашем примере – 0,0996;
- 3) **суммарное число членов когорты** для интересующего срока наблюдения;
- 4) **медиана времени дожития** (median survival time) – период времени, в течение которого изучаемый исход наступит у 50% участников исследования. Указанная величина соответствует **медиане времени до наступления исхода** (median time to event) в случае выбора альтернативного представления графика Каплана-Мейера.

В нашем примере медиану времени дожития указать невозможно, поскольку на момент окончания наблюдения более половины членов когорты не имели изучаемого исхода. Тем не менее, можно указать первый квартиль (25-й процентиль) кривой дожития – он равен **1,57 лет**; к этому сроку изучаемое состояние разовьется у 25% испытуемых лиц.

Анализируя данные когортных исследований, **необходимо помнить о возможности смешивания эффектов (confounding)**. Указанное явление наблюдается в случае, если некий фактор А, учет и анализ которого в исследовании не производился, влияет как на развитие изучаемого исхода И, так и на проявление одного или нескольких факторов (Б, В, Г), учитываемых в данном исследовании. Внешне это выглядит, как если бы факторы Б, В, Г влияли на развитие исхода И, хотя на самом деле они с ним непосредственно не связаны.

Например, употребление парентеральных наркотиков увеличивает риск смерти; при этом может иметь место смешивание эффектов – наркоманы почти всегда заболевают хроническим вирусным гепатитом С (*неучтенный фактор*), который тоже увеличивает риск смерти. Факт наркомании как таковой может увеличивать риск смерти индивидуума, но наличие неучтенного фактора (УВГС) приводит к завышению степени данного

Возможность смешивания эффектов нельзя полностью учесть на этапе планирования исследования, поскольку невозможно учесть вообще все факторы, могущие представлять какую-либо значимость. До некоторой степени возможное смешивание эффектов можно нейтрализовать на этапе анализа данных исследования посредством проведения стратификации.

Например, мы анализируем сравнительную эффективность схем антиретровирусной терапии А и В. Каждой из схем было пролечено по 1000 ВИЧ-инфицированных; отсутствие вирусологического ответа было отмечено у 300 человек, получавших схему А (30%), и у 450 человек, получавших схему В (45%), относительный риск неудачи лечения составил в данном случае $30 \div 45 = 0,67$. Можно ли вести речь о том, что схема лечения А достоверно более эффективна, чем терапевтическая схема В?

На самом деле в обеих группах имеется неучтенный фактор – количество лиц, впервые получающих антиретровирусную терапию (т.е. «АРТ-наивных»). Известно, что у лиц, ранее получавших АРТ, эффективность последующих схем существенно снижается с каждой очередной сменой терапии. Ввиду этого, учет указанного фактора критически важен для

Оказывается, в группе, пролеченной по схеме А, АРТ-наивных лиц было 800 (80%), а в группе, пролеченной по схеме В – 200 (20%), т.е. при формировании исследуемых групп рандомизация по этому признаку не проводилась. Для «подгонки» (adjustment) результатов статистических выкладок по данному фактору необходимо обратиться к базе данных исследования и выяснить частоту вирусологической неудачи лечения отдельно для АРТ-наивных лиц и больных, получавших АРТ более одного раза, для каждой из двух

Значение фактора стратификации	Схема А		Схема В		Относительный риск
	Всего	Неудача лечения (%)	Всего	Неудача лечения (%)	
АРТ-наивные	800	200 (25%)	200	50 (25%)	$25 \div 25 = 1,0$
Ранее уже получали АРТ	200	100 (50%)	800	400 (50%)	$50 \div 50 = 1,0$
Всего	1000	300 (30%)	1000	450 (45%)	$30 \div 45 = 0,67$

В обеих группах пропорция лиц с вирусологической неудачей лечения одинакова как среди АРТ-наивных лиц, так и среди ВИЧ-инфицированных, ранее уже получавших АРТ; в обоих случаях относительный риск равен 1,0. В итоге приходится признать, что на самом деле схемы лечения А и В обладают одинаковой эффективностью, а полученная в исследовании разница обусловлена только лишь ошибкой при формировании изучаемых групп, которые не были рандомизированы по фактору, оказывающему существенное влияние на результат интересующего нас вмешательства. **Важно помнить, что при стратификации уменьшается размер сравниваемых групп, что ведет к**

В качестве **типичного примера когортного дизайна** можно привести исследование, выполненное на когорте больных гемофилией Бесплатного Королевского Госпиталя (Royal Free Hospital). В гематологическом отделении вышеупомянутой больницы наблюдалось 111 ВИЧ-инфицированных мужчин-гемофиликов, заразившихся в период с 1979 по 1985 гг., на предмет особенностей течения ВИЧ-инфекции у данной категории больных. Продолжительность наблюдения составила 25 лет. Учитывались демографические факторы, клиническая симптоматика, лабораторные данные, а также информация об эффективности лечения. На момент окончания исследования в живых остались 39 его участников, в том числе 28 – под непосредственным наблюдением в составе когорты.

Другим примером «классического» когортного исследования является мультицентровое исследование когорты лиц со СПИДом. В течение двух вербовочных периодов (с апреля 1984 г. по март 1985 г. и с 1987 г. по 1991 г.) было рекрутировано несколько сотен ВИЧ-инфицированных гомосексуалистов, проживающих в 4 столичных областях США; в дальнейшем они проходили амбулаторное обследование каждые 6 месяцев. В процессе обследования собиралась демографическая информация (дата рождения, национальность), клинические данные (обследование на наличие признаков СПИДа при каждом осмотре, регистрация даты и причины смерти), лабораторные данные (подсчет уровня CD4+ лимфоцитов и количества копий РНК ВИЧ в плазме крови), сведения о проводимой терапии (даты начала и прекращения назначения всех антиретровирусных препаратов), а также некоторые другие данные (дата первого положительного обследования на ВИЧ-инфекцию, проводилась ли профилактика пневмоцистной пневмонии и

Еще одним образцом когортного дизайна является исследование частоты коинфекции ВИЧ и туберкулеза, проведенное в одном из госпиталей Нью-Йорка. Первоначально 513 внутривенных наркоманов были обследованы на ВИЧ-инфекцию; 215 из них оказались ВИЧ-положительными. Затем вся указанная когорта наблюдалась в течение 2 лет, при этом регистрировались любые признаки активного туберкулеза. Были получены следующие результаты: среди серопозитивных лиц (215) туберкулез развился у 8, среди серонегативных (298) – ни у одного из наблюдавшихся наркоманов, включенных в когорту. Риск развития туберкулеза в группе ВИЧ-положительных лиц был оценен как 0,037 (в группе ВИЧ-негативных лиц риск развития туберкулеза оказался равен нулю).

Когортный дизайн имеет следующие преимущества:

1. Можно оценить временную взаимосвязь между воздействием какого-либо фактора и развитием исхода (например, заболевания), поскольку *всегда известен порядок следования регистрируемых событий*;
2. На основании анализа данных, полученных в когортных исследованиях, можно попытаться установить причинно-следственные взаимосвязи между воздействиями и исходами (хотя *рандомизированные контролируемые исследования* в этом смысле предпочтительнее).

Недостатки когортного дизайна следующие:

1. Если исследуемый исход относится к редким явлениям, то размер когорты должен быть очень большим, а период наблюдения за ней – весьма продолжительным, что приводит к удорожанию всего исследования;

2. Когортные исследования занимают много времени и требуют вложения значительно бóльших средств, чем ранее описанные типы дизайна, что накладывает определенные ограничения как на максимальные размеры когорт, так и на продолжительность исследования; это особенно актуально для стран с невысоким уровнем ассигнований на науку;

3. Высокая вероятность систематических ошибок (biases) при сборе и анализе данных из-за «смешивания эффектов» (confounding) – ведь невозможно заранее знать, какие факторы являются существенными для развития интересующего исследователей исхода, и включить их регистрацию в план исследования; соответственно, оставшийся неучтенным фактор (факторы) может как обуславливать развитие изучаемого явления, так и взаимодействовать с регистрируемыми воздействиями, что приводит к эффектам, описанным выше;

4. Возможны проблемы с уменьшением статистической надежности исследования ввиду неизбежного уменьшения размера когорт в ходе исследования по неустановленным причинам (т.н. loss to follow-up). Поскольку когортные исследования длятся годами и десятилетиями, члены когорты могут сменить место жительства, семейный статус, умереть, а также отказаться от наблюдения и выйти из исследования, не уточняя причин.

При этом *всякий раз желательно проверять, не связан ли выход конкретного лица из научного проекта с развитием интересующего исследователей исхода*; особенно данное положение касается исследований т.н.

«психотравмирующих» (scaring) заболеваний, в частности, ВИЧ-инфекции, парентеральных гепатитов и заболеваний, передающихся половым путем (ЗППП).