Санкт-Петербургский государственный университет

**Кафедра теории управления**

**Мурашко Андрей Юрьевич**

**Магистерская диссертация**

**Оценка сбалансированности блоковой рандомизационной последовательности со стратификацией при планировании эксперимента в клинических исследованиях**

Направление 01.04.02

Прикладная математика и информатика

Методы прикладной математики и информатики в задачах управления

Научный руководитель,

доктор физ.-мат. наук,

доцент Зубов А.В.

Рецензент,   
доктор физ.-мат. наук,

доцент Михеев С.Е.

Санкт-Петербург

2018

Содержание

[Содержание 2](#_Toc514954788)

[Введение 3](#_Toc514954789)

[Постановка задачи 5](#_Toc514954790)

[Обзор литературы 6](#_Toc514954791)

[Глава 1. Клинические исследования 7](#_Toc514954792)

[1.1 Определение клинических исследований 7](#_Toc514954793)

[1.2 Виды и фазы КИ 9](#_Toc514954794)

[1.3 Рандомизация в КИ 15](#_Toc514954795)

[Глава 2. Поиск дисбаланса 19](#_Toc514954796)

[2.1 Теоретический поиск дисбаланса 19](#_Toc514954797)

[2.2 Поиск дисбаланса на практике 23](#_Toc514954798)

[Выводы 32](#_Toc514954799)

[Заключение 33](#_Toc514954800)

[Список литературы 34](#_Toc514954801)

[Приложения 36](#_Toc514954802)

[Приложение 1 36](#_Toc514954803)

[Приложение 2 40](#_Toc514954804)

# Введение

Обычно, в клинических исследованиях пациенты случайным образом назначаются в контрольные группы терапии равного размера. Контрольные группы равного размера планируются как по статистическим, так и по нормативным, эстетическим причинам. Статистический анализ наиболее эффективен, когда размеры группы равны, мощность исследования максимальна. Выбор равных групп также согласуется с предположением в начале исследования, что каждое из назначений лечения в равной степени может принести пользу пациенту, а также, что каждое из назначений лечения равновероятно может достаться участнику исследования. Предпочтительно поддерживать равные группы применения лечения в течение всего исследования. Если период регистрации достаточно длинный, чтобы произошли изменения в потенциальных переменных смешивания, таких как новые вспомогательные процедуры или новые критерии направления, поддержание равных групп в ходе исследования уменьшает потенциальный источник предвзятости регуляторными органами. В испытаниях, в которых пациентов вводят последовательно, используется модификация строгих случайных назначений, чтобы гарантировать, что группы будут иметь примерно одинаковый размер и будут уравновешиваться в течении всего времени исследования.

Один из способов достижения баланса в числе пациентов, назначенных для каждого лечения во времени, заключается в планировании назначений с использованием блоковой рандомизации [16]. В каждом блоке содержится равное количество присвоений к группе, которые перечислены в произвольном порядке. Назначения лечения участникам выдаются последовательно в блоках. По мере заполнения каждого блока, количество пациентов во всех группах лечения возвращается в идеальный баланс. В идеале группы терапии сходны по отношению к прогностическим факторам, которые могут повлиять на ход лечения. Исходя из того, что важные различия между группами могут произойти случайно, схема рандомизации может быть стратифицирована для важных ковариат, чтобы обеспечить в них баланс по группам терапии. Если используется стратификация, блоковая рандомизация может быть выполнена в пределах страт, чтобы обеспечить равные группы в каждой страте. При блоковой рандомизации со стратификацией дисбаланс все еще может иметь место для испытания в целом, особенно если имеется большое количество страт или размеры блока слишком велики по сравнению с количеством пациентов и групп лечения. При планировании должен быть достигнут некоторый компромисс между желанием стратифицировать все важные ковариаты и желанием иметь равные группы лечения как внутри слоев, так и в целом.

В диссертации рассматриваются известные подходы для поиска дисбаланса в момент старта исследования, дорабатываются, а также создается новый подход к решению данной проблемы.

Актуальность рассматриваемой темы заключается в том, что на старте клинического исследования очень часто не уделяется внимание прогнозам распределения участников по группам лечения. Из-за этого в конце исследования получается не равное распределение пациентов по группам терапии, нарушение правовых документов, таких как информированное согласие, в котором сказано, что участник может попасть в каждую из групп лечения равновероятно. А также в худшем случае невозможность правильно анализировать полученные данные, что приведет к повторению того же самого исследования с уже оцененными прогнозами дисбаланса и измененным количеством прогностических факторов. А это повторение огромных затрат бюджетов.

# Постановка задачи

При планировании рандомизационного клинического исследования со стратификацией необходимо учитывать прогнозы для дисбаланса количества участников исследования между группами терапии. А также вычислить потенциал рандомизационного списка для невозможности статистического анализа получившихся данных.

Задачи магистерской диссертации:

1. Поиск и рассмотрение существующих решений данной проблемы
2. Исследование путей вычисления оценки возможности проведения исследования с входными параметрами в виде количества участников и количества страт;
3. Практическая проверка полученных теоретических вычислений;
4. Создание программного продукта для оценки рандомизационных списков.

# Обзор литературы

* Hallstrom A, Davis K. Imbalance in treatment assignments in stratified blocked randomization // Control Clin Trials, 1988. Vol 9 No 4, P. 375–382.

В данной статье рассматривается подход к поиску дисбаланса на этапе планирования клинического исследования со стратификацией. Приводятся примеры реальных исследований и расчет на их данных относительно количества ковариат. Показаны случаи «плохих» исходов.

* Lachin J. Properties of simple randomization in clinical trials // Control Clin Trials, 1988. Vol 9 No 4, P. 312–326.

В статье представлены свойства рандомизации и правила случайного распределения. Последнее в основном используется в случаях, когда общий размер выборки n известен априори. Показана модель «Efron» для случайного смещения, она используется для оценки потенциала смещения в оценке лечебного эффекта из-за ковариационного дисбаланса. Сделан вывод о том, что полная рандомизация заслуживает больших клинических испытаний.

* Fisher R. The design of experiments. New York: Hafner 1935. 250 p.

В книге описаны основные этапы планирования практических экспериментов. Приведены обоснования каждого из этапов. Рассмотрены примеры. Особое внимание уделяется основным принципам рандомизации и контроля проведения испытаний. Рассмотрены возможные варианты распределения выборки.

* Rosenberger W, Lachin J. Randomization in Clinical Trials : theory and practice. New York: Wiley, 2002. 288 p.

Здесь приведены практически все существующие методы рандомизации участников клинического исследования. Дискуссии о существующих философиях, противоречиях и новых методов рандомизации в клинических испытаниях. Методы минимазации затрат с помощью уменьшения выборки. Авторы привлекают внимание ко всевозможным проблемам и решениям составления рандомизационных списков.

# Глава 1. Клинические исследования

**1.1 Определение клинических исследований**

**Клиническое исследование (**КИ) — научное исследование при участии людей, проводимое для того, чтобы оценить эффективность и безопасность нового лекарственного препарата (ЛП) или расширить показания к применению уже известного ЛП, а также с целью изучения безопасности и эффективности новых инвазивных (в т.ч. хирургических) и неинвазивных методов лечения и диагностирования. Проведение КИ является обязательным условием создания и вывода на рынок нового лекарственного средства, а также получения одобрения уполномоченных органов.

**Цели клинических исследований:**

* вывести новый фармацевтический препарат на рынок;
* оценить терапевтическую или профилактическую эффективность и переносимость нового лекарственного средства;
* установить наиболее рациональные дозы и разработать схемы их применения;
* сравнить с уже существующими препаратами.

Во всем мире принятым является выполнение требований международного стандарта надлежащей клинической практики (Good Clinical Practice; GCP) [8], [9] для проведения всех клинических испытаний лекарственных средств (ЛС) или терапий. В Российской Федерации необходимо руководствоваться национальным стандартом ГОСТ Р 52379-2005 «Надлежащая клиническая практика» - аналогом GCP в нашей стране [17].

Основные элементы GCP:

* защита прав субъекта исследования - человека;
* информированное согласие;
* этический комитет: спонсор – монитор – исследователь;
* стандартизированные рабочие инструкции по проведению исследования и отчетности;
* оценка эффективности и безопасности;
* требование достоверности;
* видимость всех побочных/нежелательных явлений.

Обычно в клиническом исследовании принимают участие несколько ключевых сторон: спонсор, контрактно-исследовательская организация (КИО), исследователи и пациенты. Иногда роль КИО берет на себя сам спонсор, но такая практика не всегда признается регуляторным органом, принимающим отчет. При планировании клинического исследования первым делом представители компании-спонсора (или же КИО по поручению спонсора) определяют необходимый размер выборки пациентов (количество участников исследования, которых необходимо привлечь), чтобы получить статистически значимый результат, который покажет разницу в эффективности и безопасности сравниваемых препаратов (терапий). Количество участников исследования зависит от фазы, изучаемых параметров, заболевания и т.д.

В основном, сами фармацевтические компании спонсируют и проводят клинические исследования. Так же проекты могут финансировать академические учреждения (институты) и исследовательские центры. В структуре мирового рынка клинических исследований доля КИО в денежном выражении составляет примерно 25%. Для стран Центральной и Восточной Европы и СНГ этот показатель оценивается на уровне 50–60%. Чтобы провести исследование, КИО и фармацевтические компании заключают договор с мониторами клинических исследований (clinical research associate)— квалифицированным персоналом, ответственным за мониторинг процесса и сотрудничество с исследователями.

Согласно данным аналитической компании «PricewaterhouseCoopers», мировой рынок клинических исследований, оценивается на уровне 50–80млрд дол. США. По информации, опубликованной на *ClinicalTrials.gov*, созданном Национальным институтом здоровья США (National Institutes of Health — NIH) при сотрудничестве с Управлением по контролю за пищевыми продуктами и лекарственными средствами США (Food and Drug Administration — FDA) во исполнение «FDA Modernization Act», по итогам 2011 года во всем мире зарегистрировано более 18 тыс. клинических исследований». Наибольшее количество клинических исследований в регио­нальном разрезе проводится в Северной Америке, на 2-й позиции находится Европа, а третье место у Восточной Азии. Ключевыми рынками последней стали Китай (677 проектов) и Южная Корея (774 проекта).

Согласно данным американской Ассоциации разработчиков и производителей лекарственных препаратов (PhRMA), только лишь 250 лекарств из 10000 лекарств-кандидатов, поступивших в обработку фармацевтическими компаниями США, выходят на этап КИ. На стадию клинических исследований попадают из них лишь 5. В общую медицинскую практику и на рынок поступает лишь один кандидат – он становится лекарственным средством.

Разрешение на маркетинг получают в США только лишь 11% препаратов для лечения нарушений обмена веществ, 14 % препаратов для лечения заболеваний ЦНС, 15% кардиологических препаратов, 20% препаратов для лечения заболеваний дыхательной системы, 27% онкологических препаратов и 40% антибиотиков, перешедших на этап клинических исследований. Только 34 препарата американское Управление по контролю над пищевыми продуктами и лекарственными препаратами (FDA) зарегистрировало в 2009 году, 31 - в 2008 году, 23 - в 2007 году.

**1.2 Виды и фазы КИ**

Для чистоты эксперимента на практике очень часто не разглашается принадлежность пациента к одной из групп лечения. В отношении пациента это происходит для получения более достоверных данных, например, чтобы был возможен «эффект плацебо». В отношении исследователя\врача для невозможности подлога данных при распределении пациентов. Поэтому КИ бывает нескольких видов:

**Открытое исследование** предполагает, что обе стороны: и пациент, и врач, после проведения рандомизации узнают вид лечения, который будет применен. Такой вид исследований обычно используют в тех случаях, когда очень сложно или этически нецелесообразно создать полную имитацию исследуемого вмешательства, или введение лекарства носит ярко выраженные признаки, которые невозможно скрыть. Например, когда сравнивают новую форму приема ЛС, со старой.

**Простое слепое исследование** заключается в том, что пациенту до окончания исследования не сообщается о виде применяемого. Данный аспект документируется в информированном согласии. После проведения рандомизации врачу станет известно о виде лечения пациента.

**Двойное слепое исследование** — ни исследователю, ни пациенту не разглашается информации о группе применяемой. Перед началом исследования необходимые препараты подготавливают к введению на клинической базе и нумеруют, согласно созданному медицинскому листу. Как минимум один человек из команды исследователей (непосредственно занимающийся приготовлением раствора и нумерацией) будет точно знать, что именно получает пациент. Обычно, в таких случаях на клиническую базу делегируются два независимых монитора клинических исследований, один из которых проверяет документацию только «ослепленых» исследователей, а второй — только «неослепленных». Чрезвычайно важно, что неразглашение распределения больных по группам лечения принципиально для непредвзятой оценки и самим пациентом, и наблюдающим за ним «ослепленным» исследователем результатов лечения, поэтому «неослепленный» персонал ни при каких обстоятельствах не должен делиться имеющейся у него информацией со своими коллегами.

**Тройное слепое исследование** — о типе вмешательства в каждой из сравниваемых групп не знают ни пациент, ни исследователь, ни статистик, обрабатывающий результаты исследования. В ключевых исследованиях часто создается независимый комитет по оценке данных, который проводит промежуточные анализы эффективности и безопасности до полного укомплектования и закрытия базы данных. Конечно же, членам комитета важно получить уже расшифрованные данные, то есть знать, какое лечение получал каждый из пациентов. В случае тройного слепого исследования для подготовки отчетов комитету используется отдельный независимый «неослепленный» статистик или группа статистиков.

Кроме доклинических испытаний — исследований препаратов на животных, а также построение приближенной модели, для вывода нового лекарственного средства на рынок оно должно пройти как минимум 3 фазы исследования на людях.

**Фазы клинических исследований:**

**Фаза I** - клинико-фармакологическая

Служит для изучения фармакологических свойств и подтверждения профиля безопасности нового ЛС у здоровых участников исследования. Главная задача - оценить переносимость исследуемого препарата, токсичность, установить наличие у него терапевтического действия, определить параметры фармакокинетики, создать рациональную основу для выбора доз, схем и форм применения. Исследования обычно проводят при участии небольшого числа здоровых добровольцев (обычно до 50).

**Фаза II** — контролируемые или пилотные исследования

Целью фазы II является установка эффективности и безопасности ЛС на определенном контингенте пациентов, включающем 100-200 человек, и разработка оптимальных режимов дозирования.

Фаза IIa - пилотное исследование - проводится для поиска дополнительных фармакологических свойств ЛС у пациентов.

Фаза IIb - контролируемое исследование - предусматривает наличие контроля или контрольной группы, для недопущения таких неточностей, которые связанны с влиянием различных независимых или прогностических факторов на результат терапии. Желательно, чтобы основная и контрольная группы не различались по возрасту, полу, тяжести заболевания и другим факторам, это возможно достичь, используя рандомизацию.

Используется четыре типа контроля:

1. Исходного состояния

Контроль исходного состояния всегда используется при проведении любых КИ. Измерения на скрининге производятся для каждого пациента для получения исходных данных, которые будут сравнивать с результатами, полученными во время и после окончания исследования.

1. Плацебо

Способ плацебо-контроля или «негативного» контроля заключается в применении участником исследования некоторого плацебо - неактивного вещества, не отличимого ни по каким признакам (внешний вид, вкус, запах) от исследуемого препарата.

1. Активный

В активный (позитивный) контроль включается лечение с применением ЛС, которое является эффективным для исследуемого показания (ЛС активного контроля). Как и в случае применения плацебо, ЛС активного контроля не отличается от изучаемого препарата.

4. По архивной статистике

Контроль по архивной статистике или исторический контроль позволяет сравнить исследуемую терапию с существующими данными о вариантах исходов конкретного заболевания. Данный способ контроля используется тогда, когда не существует других вариантов эффективного метода лечения исследуемой патологии или редкого заболевания.

**Фаза III** - расширенные клинические исследования

Главной целью расширенных клинических исследований является получение дополнительной информации о безопасности и эффективности новых препаратов у пациентов в условиях, максимально приближенных к клинической практике. В ходе этих исследований изучаются особенности действия ЛС у больных с сопутствующими заболеваниями таких, как нарушение функции печени и почек кровообращения. Оцениваются терапевтические преимущества - влияние на уровень жизни в целом. Выявляются редко встречающиеся побочные реакции (нежелательные явления) и особенности взаимодействия нового препарата с другими ЛС, а также изучаются фармакоэкономические аспекты. Фаза III клинических исследований завершается представлением препарата на регистрацию, в связи с чем она делится на две фазы:

• фаза IIIa – это период до представления заявки на регистрацию;

• фаза IIIb - это период с момента подачи заявки на регистрацию и до окончательной регистрации препарата в контролирующих органах.

**Фаза IV** - пострегистрационные исследования

Проведение IV фазы возможно только после разрешения применения нового препарата в медицинской практике. Цель этой фазы - изучение перспективы расширения показаний к применению ЛС, усовершенствование режимов назначения и схем лечения, а также длительное наблюдение (в течение многих лет) для рекламы. Отдельное внимание обращают на сбор и анализ информации о нежелательных явлениях новых ЛС. Еще более важным представляется изучение отдаленной эффективности препаратов на выживаемость, которое проводится у многих сотен и тысяч больных.

Контролируемые исследования фазы IV могут включать от нескольких десятков до тысяч пациентов и бывают ретроспективные и проспективные. Ретроспективные исследования проводятся на основе прошлого опыта применения разных ЛС или видов терапии по данным историй болезни. Проспективные исследования планируются на перспективу (до начала набора больных) и проводятся по общему протоколу в сбалансированных группах больных, что значительно повышает надежность полученных результатов.

**1.3 Рандомизация в КИ**

**Мощность критерия** относительно определенной альтернативы — вероятность правильного выбора нулевой гипотезы, или ее отклонения в случае, когда альтернативная гипотеза действительно истинна.

**Статистическая мощность** бинарной гипотезы — вероятность того, что правильно отвергается нулевая гипотеза, когда альтернативная гипотеза верна. Чем выше статистическая мощность, тем меньше вероятность совершить ошибку второго рода (ложноотрицательное событие).

**Мощность теста (мощность исследования)** — характеристика исследования, которая позволяет выявить клинически важные различия между исследуемым препаратом и препаратом сравнения, если таковые существуют в действительности. Чем больше выборка пациентов, тем больше мощность теста или мощность исследования. Но чем больше количество пациентов, тем дороже будет стоить проведения клинического исследования. Поэтому нужно всеми доступными методами оптимизировать количество участников, не уменьшая запланированной мощности исследования. Мощность исследования достигает максимума при распределении по группам терапии в соотношении один к одному.

Рандомизация – это ключевое понятие анализа медицинских данных. Этим термином обозначают способ случайного генерирования выборки. Когда говорят, что группа данного размера является простой случайной выборкой из большей группы, подразумевается, что все возможные выборки этого размера извлекают с равными вероятностями. Выражение «обработка назначается объектам случайно» обозначает, что вероятность назначения каждого вида обработки одинакова для всех объектов. На необходимость рандомизации впервые указал Р. Фишер [6]. Основными целями рандомизации являются: исключение возможности совершения исследователем преднамеренной или невольной ошибки, соблюдение протокола исследования, а также сохранение заявленной статистической мощности.

Например, необходимо провести клинические исследование ЛС, чтобы установить его эффективность. Для этого больных делят пополам на группы, одним назначается лекарство, другим – нейтральный препарат. Испытуемые поступают на испытания сериями, не одновременно, в течение какого-то времени.

Применим два вида рандомизации

1. В интервале от 1 до 100 выберем 50 различных чисел, активное лекарство будет назначено тем больным, чьи номера попали в эту выборку. Другие 50 пациентов получат нейтральный препарат.

Можно найти два недостатка у такого метода:

Во-первых, если исследование по какой-либо причине завершится раньше, то общее число пациентов, принимавших «пустышку», с большой вероятностью не будет равно числу пациентов, принимавших активное ЛС.

Во-вторых, если клиническое состояние пациентов, включающихся в испытание в один момент времени, отличается от состояния пациентов, включающихся в другой момент, или меняются правила приема препаратов, то, несмотря на рандомизацию, две группы, возможно, будут отличаться по типу пациентов или по правилам приема лекарств.

Метод блоковой рандомизации сохраняет одинаковую численность всех групп в течение всего процесса рандомизации.

Больных, которых планируется включить в исследование, условно разделяют на равные части (блоки). В пределах блока методы лечения распределяют так, чтобы разными методами лечилось одинаковое число пациентов, но последовательность назначения лечения была бы различной. Затем блоки случайно распределяют, например, пользуясь таблицей случайных чисел.

С помощью этого метода легко провести независимую последовательную рандомизацию по группам пациентов, поступающих в течение коротких промежутков времени.

Пусть ежемесячно в испытаниях начинают участвовать десять больных. Разумно случайно назначать пяти пациентам лечение одного вида, а остальным пяти пациентам — другого, повторяя случайное назначение каждый месяц, по мере поступления новых партий больных.

Частный случай этого метода — испытания на парах пациентов, когда один из двух пациентов получает активный, а другой - нейтральный препарат.

В некоторых случаях необходимо равномерно разделить пациентов для дополнительных сравнений по некоторым **прогностическим факторам** —характеристики участников исследования, для которых может быть более высокая или более низкая вероятность благоприятных или неблагоприятных исходов КИ. Это позволяет соблюдать баланс важных характеристик между разными группами терапии.

В таких случаях используется **стратификация** — метод формирования выборки, при котором совокупность всех участников, соответствующих критериям включения в исследование, сначала разделяется на страты на основе одного или нескольких факторов (например, пола, возраста, и т.п.), потенциально влияющих на изучаемый исход, а затем в каждой из этих страт независимо проводится набор участников в группы лечения [12].

Обо всех возможностях и вариантах рандомизации написано в [14].

Рассмотрим пример блокового рандомизационного списка со стратификацией:

Таблица 1 - рандомизационный список

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Num** | **Strata** | **Group** | **Block** |  | **Num** | **Strata** | **Group** | **Block** |
| R001 | Женщина | Group\_1 | 1 | R021 | Мужчина | Group\_2 | 1 |
| R002 | Женщина | Group\_1 | 1 | R022 | Мужчина | Group\_1 | 1 |
| R003 | Женщина | Group\_2 | 1 | R023 | Мужчина | Group\_1 | 1 |
| R004 | Женщина | Group\_2 | 1 | R024 | Мужчина | Group\_2 | 1 |
| R005 | Женщина | Group\_1 | 2 | R025 | Мужчина | Group\_1 | 2 |
| R006 | Женщина | Group\_2 | 2 | R026 | Мужчина | Group\_2 | 2 |
| R007 | Женщина | Group\_1 | 2 | R027 | Мужчина | Group\_1 | 2 |
| R008 | Женщина | Group\_2 | 2 | R028 | Мужчина | Group\_2 | 2 |
| R009 | Женщина | Group\_2 | 3 | R029 | Мужчина | Group\_1 | 3 |
| R010 | Женщина | Group\_1 | 3 | R030 | Мужчина | Group\_2 | 3 |
| R011 | Женщина | Group\_1 | 3 | R031 | Мужчина | Group\_2 | 3 |
| R012 | Женщина | Group\_2 | 3 | R032 | Мужчина | Group\_1 | 3 |
| R013 | Женщина | Group\_2 | 4 | R033 | Мужчина | Group\_2 | 4 |
| R014 | Женщина | Group\_2 | 4 | R034 | Мужчина | Group\_1 | 4 |
| R015 | Женщина | Group\_1 | 4 | R035 | Мужчина | Group\_1 | 4 |
| R016 | Женщина | Group\_1 | 4 | R036 | Мужчина | Group\_2 | 4 |
| R017 | Женщина | Group\_2 | 5 | R037 | Мужчина | Group\_1 | 5 |
| R018 | Женщина | Group\_1 | 5 | R038 | Мужчина | Group\_2 | 5 |
| R019 | Женщина | Group\_2 | 5 | R039 | Мужчина | Group\_1 | 5 |
| R020 | Женщина | Group\_1 | 5 | R040 | Мужчина | Group\_2 | 5 |

При создании данного списка преследовалась цель равномерного распределения 20 участников исследования по 2 группам терапии, учитывая один прогностический фактор – пол. Как видно из списка, создается полный рандомизационный список для каждой из страт, так как мы не можем заблаговременно сказать, сколько будет женщин или мужчин, даже придет ли хоть 1 человек в каждую из страт. Именно такой подход и может породить **дисбаланс** (разность между количеством участников в каждой из групп терапии) между группами терапии. Для примера рассмотрим случаи распределения: например, пришло 11 женщин и 9 мужчин. Тогда в группе 1 будет 11 участников, а в группе 2 только 9.

Из-за выбывания в начале исследования даже при рандомизации без стратификации среди групп терапии может произойти дисбаланс, данные исследования проводились в статьях [1], [10], [11], [13], [15].

# Глава 2. Поиск дисбаланса

**2.1 Теоретический поиск дисбаланса**

Существует бесконечное количество оценок рандомизационных списков, когда исследование уже закончено. Но на этом этапе оценивать уже поздно. Это дает лишь знания на счет будущего исследования. Лучше определиться с этим на этапе планирования.

Рассмотрим общий принцип вариантов распределения при рандомизации со стратификацией [7]:

Для простоты рассмотрим вариант клинического исследования с двумя группами терапии. Пусть N пациентов должны быть назначены с равной степенью вероятности в группы исследуемого препарата и группу контроля в k стратах. Пусть присвоения во всех стратах сбалансированы в блоках равного размера . Пусть – число участников в последнем блоке страты i,

Пусть – число участников, попавших в группу исследуемого препарата в последнем блоке,

Дисбаланс для i-ой страты равен ,

Так же обозначим формулу для максимального размера дисбаланса:

Если , тогда .

В общем случае на имеет гипергеометрическое распределение, полученное путем выборки элементов из популяции , в которой содержится участников из группы исследуемого препарата и участников из группы контроля. По формулам гипергеометрического распределения мы можем найти для математическое ожидание () и дисперсию )

Математическое ожидание

Дисперсия

,

Тогда мы можем посчитать дисперсию дисбаланса:

(1)

И тогда общая дисперсия дисбаланса равна сумме всех дисперсий на каждом уровне стратификации:

(2)

Для использования формул (1) и (2) нам необходимо знать распределение для . Пусть – число пациентов, рандомизированных в страту i. Тогда вероятность того, что пациент находится в iой страте может быть высчитана по формуле:

имеет биномиальное распределение с параметрами и N. Тогда математическое ожидание для формул (1) и (2) можно вычислить, основываясь на биномиальном распределении:

(3)

(4)

На практике вычислить значения невозможно, чтобы решить реальную прикладную задачу, необходимо ввести некоторые допущения. Для этого рассмотрим 2 возможных особых случая для упрощения данных формул (3) и (4):

1. Допустим, что ожидаемое количество пациентов в каждой страте велико, относительно размера блока, тогда мы можем предположить, что распределение R равномерно по целым числам от 1 до B. Тогда пересчитаем мат. ожидание для :

,

.

А дисперсия для дисбаланса становится равной:

(5)

Получается, что если рассматривать случай, когда во всех стратах блоки одинакового размера, то:

(6)

1. Если же количество пациентов в каждой страте мало относительно размера блока, то мы должны рассматривать биномиальное распределение с параметрами N и . Тогда перепишем формулы (3) и (4):

Для простоты записи рассмотрим только 1ое слагаемое:

Вернемся к общим вычислениям :

Тогда из (2) мы можем вывести формулу дисперсии:

(7)

Как видно из вышеперечисленных формул, размер суммарного дисбаланса прямо пропорционален размеру блока и количеству страт. Т.к. величиной, которую мы можем самостоятельно выбирать, является только размер блока, то, исходя из правил рандомизации, мы можем сказать, что для наименьшего дисбаланса наилучший размер блока должен быть подсчитан по формуле:

(8)

где Т – обозначает количество групп терапии.

В случае если на всех уровнях стратификации размеры блоков равны, тогда максимальный размер дисбаланса равен:

(9)

Таким образом, при необходимости распределить пациентов в более, чем 2 группы терапии мы можем вычислить дисперсию, воспользовавшись многомерным гипергеометрическим распределением.

**2.2 Поиск дисбаланса на практике**

Попробуем использовать выведенные формулы на реальной задаче [5]. Чтобы она была наглядной используем достаточно экстремальный случай с большим количеством страт и малым размером выборки. Рассмотрим пример исследования, в котором нам необходимо рандомизировать 64 участника в две равные группы («Группа 1» и «Группа 2»), учитывая четыре прогностических фактора у каждого из которых два уровня, всего 16 страт, о долях ожидаемого распределения (доля здесь – это процентное соотношение ожидаемого количества субъектов в одной группе к другой, она может быть обозначена как вес) по стратам информации нет, так что примем его одинаковым:

1. Мужчина/Женщина;
2. Наличие/отсутствие сахарного диабета;
3. Старше или младше 35 лет на момент подписания информированного согласия.
4. Курит/не курит

Из (8) определяем наилучший размер блока .

Из (9) .

Так как ожидаемое количество участников в каждой страте больше размера блока можем вычислить дисперсию дисбаланса по (6)

Среднеквадратичное отклонение (СО)

(10)

Теперь проверим полученные значения с помощью SAS 9.4.

Создадим 150 различных рандомизационных списков (Приложение 1).

Заполним каждый из них 1000 раз случайными комбинациями пришедших пациентов и посчитаем получившееся среднеквадратичное отклонение, коэффициенты разброса (Приложение 2). Будем считать, что с помощью такого количество вероятностей можем полностью описать возможности данного распределения.

После заполнения рандомизационных списков находим общее количество участников в каждой из групп терапии:

proc freq data=rand noprint;

table med\_grp/ out = GRFreq;

run; /\*Подсчет учасников в каждой группе терапии в каждой страте\*/

Находим дисбаланс, а также коэффициент разброса от идеального распределения:

(11)

data all\_&k;

set all\_&k;

k=k;

coef=Group\_1/Group\_2;

Disb=Group\_1-Group\_2;

run;

Так как распределение коэффициента разброса (11) нормальное, то с помощью одновыборочного t-критерия, можем проверить вероятность удаления его от 1 и наглядно посмотреть на варианты распределений. Зададим нулевую гипотезу: H0 - коэффициент равен единице. Тогда альтернативная гипотеза: H1 - коэффициент не равен единице.

proc ttest data = all\_&k h0 = **1**;

var coef;

run;/\*Проверка каждого из списков по отдельности с помощью t-теста\*/

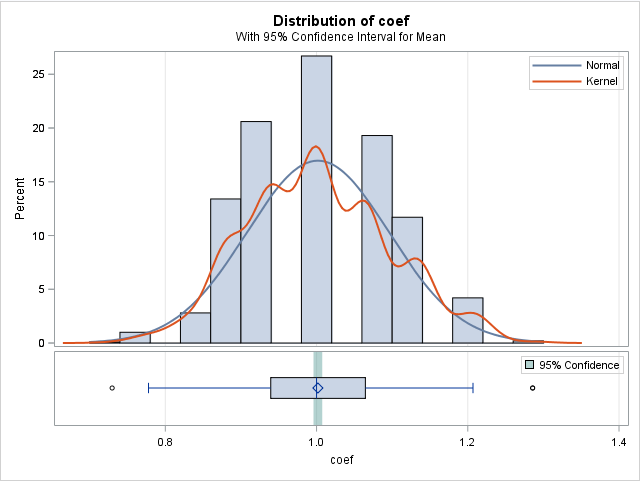
Пример 1:  


Рисунок 1 – Распределение по группам терапии

На данном графике (рис.1) представлен один из вариантов распределения пациентов по группам терапии, следовательно при среднем от коэффициента распределения меньше единицы (11) большинство участников исследования попало в группу 2, при среднем от коэффициента распределения больше единицы наоборот, в группу 1. В данном примере среднее коэффициентов распределения примерно равно единице, поэтому среднее дисбалансов будет равно нулю.

Так же SAS 9.4 выдал следующий результат о распределении коэффициентов:

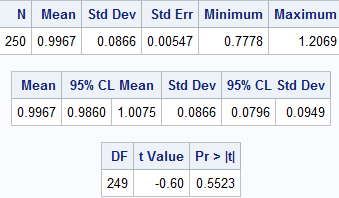


Рисунок 2 – Описательная статистика коэффициента распределения

Как видно из рисунка 2, р-значение =0,5523. Следовательно, мы можем отвергнуть альтернативную гипотезу и принять нулевую, о том, что коэффициент распределения =1.

Дополнительно для данного случая случая рассмотрим описательную статистику для дисбаланса, считаем ее с помощью:

proc means data = all\_138 noprint;

var Disb;

output out=means\_138 n = coln\_N mean = coln\_mean std=coln\_StdDev median = coln\_Median min=coln\_Min max=coln\_Max lclm=coln\_lclm uclm=coln\_uclm;

run;

Среднее значение = -0,056, СО =3,052, минимум=-10, максимум=8, медиана=0.

А также частотное распределение (таблица 2).

Таблица 2 – частотная таблица дисбаланса

|  |  |  |
| --- | --- | --- |
| Дисбаланс | Частота | Процент |
| -10 | 1 | 0.1 |
| -8 | 9 | 0.9 |
| -6 | 44 | 4.4 |
| -4 | 111 | 11.1 |
| -2 | 205 | 20.5 |
| 0 | 264 | 26.4 |
| 2 | 219 | 21.9 |
| 4 | 94 | 9.4 |
| 6 | 47 | 4.7 |
| 8 | 6 | 0.6 |

После проведения симуляции на 150 рандомизационных списков для данного экстремального примера в 20 случаях нулевая гипотеза была принята, таким образом удалось выявить 130 случаев, когда использование случайного рандомизационного списка могло привести к серьезному снижению мощности при проведении анализа клинического исследования, т.к. количество пациентов крайне мало. 20 рандомизационных списков были приняты как приемлемые к использованию в клиническом исследовании с данным дизайном. Большинство распределений показали отклонение дисбаланса от 0 в одну или другую сторону, пример отклонения распределения к группе 1 представлен ниже.

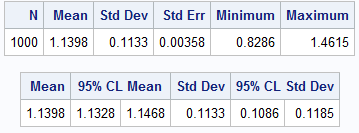


Рисунок 3 – Описательная статистика коэффициента распределения

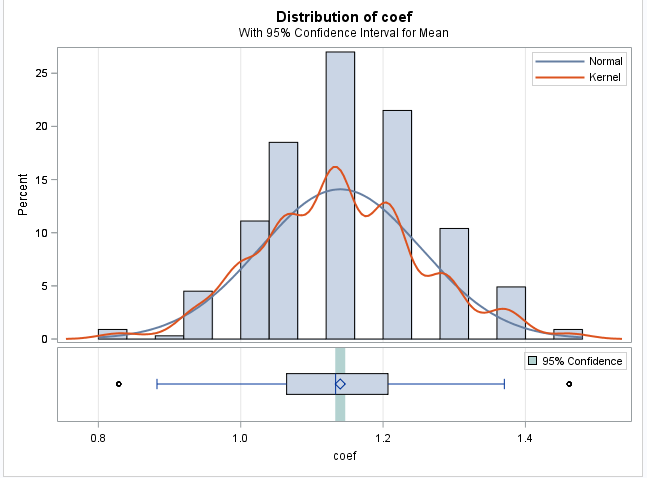


Рисунок 4 – Распределение по группам терапии

В данном случае мы принимаем гипотезу, о том, что коэффициент распределения отличен от единицы. Как видно (рис.3,4), при таком рандомизационном списке намного чаще встречается, что группа 1 больше группы 2. Посмотрим на описательную статистику для значений дисбаланса: среднее значение = 4,014, СО =3,178, минимум=-6, максимум=12, медиана=4. А также на частоты распределений (таблица 3).

Таблица 3 – частотная таблица дисбаланса

|  |  |  |
| --- | --- | --- |
| Дисбаланс | Частота | Процент |
| -6 | 1 | 0.9 |
| -4 | 9 | 0.3 |
| -2 | 44 | 4.5 |
| 0 | 111 | 11.1 |
| 2 | 205 | 18.5 |
| 4 | 264 | 27 |
| 6 | 219 | 21.5 |
| 8 | 94 | 10.4 |
| 10 | 47 | 4.9 |
| 12 | 9 | 0.9 |

Рассмотрев эти примеры (а также остальные 148 случаев) можно заметить, что количество исходов у дисбаланса обычно одинаково, так же, как и распределение по частотам относительно центрального (наибольшего по частоте) значения.

Из этого можно сделать вывод, что равномерность распределения по группам терапии напрямую зависит от самого рандомизационного списка. В нем может быть частота встречающихся значений дисбаланса либо равна нулю, либо же отклонена в большую или меньшую сторону от нуля.

Пример 3:

Рассмотрим теперь распределение в общем случае, возьмем все 150 рандомизационных списков (150000 вариантов распределений) и посчитаем статистику для коэффициента распределения и дисбаланса.

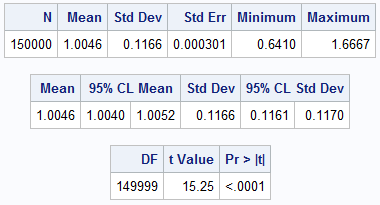


Рисунок 5 – Описательная статистика для коэффициента распределения

Рисунок 6 – Общее распределение по группам терапии

Проценты

центы

Коэффициент распределения

Проценты

центы

Коэффициент распределения

нт распределения

роцеКоэффициент распределения

0,64

нты

Коэффициент распределения

циент распределе0,64

ния

0,82

0,94

1,07

1,21

1,37

0,68

0,64

0,72

1,56

0,78

0,88

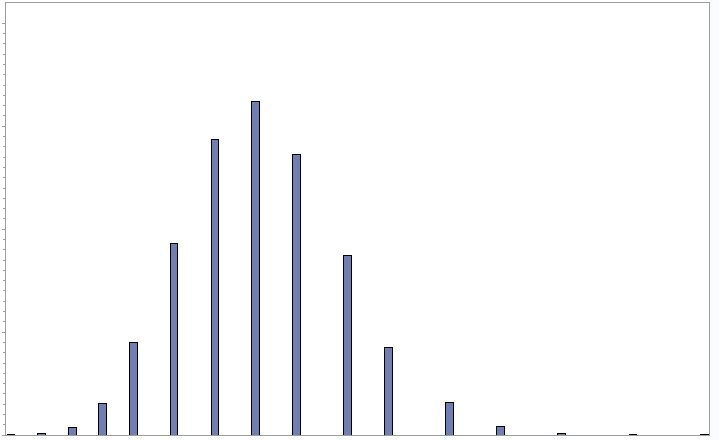
1

1,13

1,28

1,46

1,66



26

20

10

0

Из рисунков 5 и 6 видно, что общее распределение никуда не смещено. Среднее равняется единице, распределение равномерно. В нашем случае более важно посмотреть на описательную статистику дисбаланса: среднее значение = -0.06684, СО =3,681, минимум=-14, максимум=16, медиана=0.

Для наглядности также приведена таблица 4.

Таблица 4 – Общая частотная таблица дисбаланса

|  |  |  |
| --- | --- | --- |
| Дисбаланс | Частота | Процент |
| -14 | 16 | 0,01 |
| -12 | 172 | 0,11 |
| -10 | 708 | 0,47 |
| -8 | 331 | 2,08 |
| -6 | 9021 | 6,01 |
| -4 | 18607 | 12,4 |
| -2 | 28718 | 19,15 |
| 0 | 32386 | 21,59 |
| 2 | 27207 | 18,14 |
| 4 | 17449 | 11,63 |
| 6 | 8517 | 5,68 |
| 8 | 3137 | 2,09 |
| 10 | 791 | 0,53 |
| 12 | 138 | 0,09 |
| 14 | 17 | 0,01 |
| 16 | 2 | 0,00 |

Вспомним (10) разница между практически посчитанным СО и теоретически выведенным , это менее одного процента, что является статистически не значимым.

Проспективные методы исследования мощности теста далеко не полностью разработаны и приведены к оптимальным, каждая работа привносит новую информацию. При планировании исследования необходимо проверять возможность статистически рационального проведения заданного исследования.

# Выводы

Проведенная работа показывает влияние изначального определения факторов распределения на ход всего клинического исследования. Необходимо заранее проверять риски возникновения дисбаланса между группами терапии. Как было показано, это можно сделать как по приведенным формулам (5), (6), (7), так и практически. Плюсом практической проверки является то, что кроме общего поиска дисбаланса, можно посмотреть на возможные распределения участников исследования внутри уже построенного рандомизацинного списка и обновить его, в случае большой вероятности возникновения дисбаланса. Если мы знаем доли ожидаемого распределения по стратам, то результаты проверки будут абсолютно схожи с реальными данными. Все это может предотвратить огромную работу и вложение значительных денежных средств на повторное исследование.

# Заключение

В ходе выполнения магистерской диссертации были исследованы существующие методы поиска дисбаланса между группами лечения в клиническом исследовании. Были найдены новые пути решения данной проблемы, проверены на практике. Так же был разработан программный продукт с использованием языка программирования SAS 9.4 (Statistical Analysis System) для оценки рандомизационных списков.

# Список литературы

1. Altman D., Bland J. Treatment allocation by minimisation // BMJ, 2005. Vol 330 No 7495, P. 843.
2. Anisimov V. Effect of imbalance in using stratified block randomization in clinical trials // Conference: 56 Annual Session of the International Statistical Institute, 2011. Vol 10 No 1, P. 5938-5941.
3. Anisimov V. Impact of stratified randomization in clinical trials // Advances in model-oriented design and analysis, 2010. P. 1-8.
4. Anisimov V., Fedorov V. Design of multicentre clinical trials with random enrolment // In Advances in Statistical Methods for the Health Sciences, 2006. P. 387–400.
5. Database of privately and publicly funded clinical studies http://clinicaltrials.gov/
6. Fisher R. The design of experiments. New York: Hafner 1935. 250 p.
7. Hallstrom A, Davis K. Imbalance in treatment assignments in stratified blocked randomization // Control Clin Trials, 1988. Vol 9 No 4, P. 375–382.
8. ICH Topic E6, Guideline for good clinical practice E6(R1) https://www.ich.org/fileadmin/Public\_Web\_Site/ICH\_Products/Guidelines/Efficacy/E6/E6\_R1\_Guideline.pdf
9. ICH Topic E9, Statistical Principles for Clinical Trials. London 1998. http://www.ema.europa.eu/docs/en\_GB/document\_library/Scientific\_guideline/2009/09/WC500002928.pdf
10. Kundt G., Comparative evaluation of balancing properties of stratified randomization procedures // Methods Inf Med, 2009. Vol 28 No 2, P. 129-134.
11. Kundt G., Glass A. Evaluation of imbalance in stratified blocked randomization: some remarks on the range of validity of the model by Hallstrom and Davis // Methods Inf Med, 2012. Vol 51 No 1, P. 55-62.
12. Lachin J. Properties of simple randomization in clinical trials // Control Clin Trials, 1988. Vol 9 No 4, P. 312–326.
13. Matts J., Lachin J.M. Properties of permuted-block randomization in clinical trials // Control Clin Trials, 1988. Vol 9 No 4, P. 327–344.
14. Rosenberger W, Lachin J. Randomization in Clinical Trials: theory and practice. New York: Wiley, 2002. 288 p.
15. Thernau T. How many stratification factors are “too many” to use in a randomization plan? // Control Clin Trials, 1993. Vol 14 No 2, P. 98–108.
16. Zelen M. The randomization and stratification of patients to clinical trials // Journal of Chronic Diseases, 1974. Vol 27 Issue 7-8, P. 365-375.
17. ГОСТ Р.52379-2005 Надлежащая клиническая практика http://ctppcp.su/docs/iec9.pdf

# Приложения

**Приложение 1**

**%Macro** rand (siten, bl, tr, trt\_num);

%let iter=%eval(&treat./&trt\_num.); /\*Количество участников одной группы в одном блоке\*/

data rannum(keep=seed);

do i=**1** to &siten;

rannum=ranuni(&seed);

seed=int(rannum\***1000000**);

output;

end;

run; /\*Генерируем датасет из случайных чисел, объема количества страт\*/

proc sql ;

select seed

into :seed1 - :seed&siten.

from rannum;

quit; /\*Присвоение значений из rannum макропеременной seed1-seedn\*/

data rnd;

set \_null\_;

run; /\*Создание пустого датасета\*/

%do j=**1** %to &siten;

proc plan seed=&&seed&j;

factors block=&bl. ordered treat=&tr. random/noprint;

output out=data&j

treat nvals=(

%do k=**1** %to &trt\_num;

%do l=**1** %to &iter;

&k %end;

%end;

) random;

run; /\*Создание рандомизационного списка для каждой из страт\*/

data data&j.(keep= block treat Site);

set data&j;

site=&j;

run;

data rnd;

set rnd data&j;

run; /\*Склеивание ранд. списков для каждой страты в один\*/

%end;

**%mend**;

/\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*/

/\*\*\*\* Создание рандомизационного списка \*\*\*\*\*\*\*\*/

/\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*/

**%macro** rnd(seed=);

%do i=**1** %to **150**; /\*Задаем количество выходных списков\*/

%let seed=&seed+&i; /\*Случайное число для генерации списка\*/

%let siten=16; /\*Количество страт\*/

%let block=16; /\*Количество блоков\*/

%let treat=4; /\*Количество участников в группе\*/

%let trt\_num=2; /\*Количество групп терапии\*/

Proc format;

value treat

**1**='Group\_1'

**2**='Group\_2';

run; /\*Названия групп терапии\*/

%***rand*** (&siten, &block, &treat, &trt\_num);

data rnd\_&i lib.rnd\_&i;

length Strata $**60**;

set rnd;

pid='R'||put(\_n\_, z3.);

med\_grp=put(treat,treat.);

if site=**1** then Strata="Наличие сахарного диабета, женщина, младше 35 лет, курит";

else if site=**2** then Strata="Отсутствие сахарного диабета, мужчина, младше 35 лет, курит";

else if site=**3** then Strata="Наличие сахарного диабета, мужчина, младше 35 лет, курит";

else if site=**4** then Strata="Отсутствие сахарного диабета, женщина, младше 35 лет, курит";

else if site=**5** then Strata="Наличие сахарного диабета, женщина, старше 35 лет, курит";

else if site=**6** then Strata="Отсутствие сахарного диабета, мужчина, старше 35 лет, курит";

else if site=**7** then Strata="Наличие сахарного диабета, мужчина, старше 35 лет, курит";

else if site=**8** then Strata="Отсутствие сахарного диабета, женщина, старше 35 лет, курит";

else if site=**9** then Strata="Наличие сахарного диабета, женщина, младше 35 лет, не курит";

else if site=**10** then Strata="Отсутствие сахарного диабета, мужчина, младше 35 лет, не курит";

else if site=**11** then Strata="Наличие сахарного диабета, мужчина, младше 35 лет, не курит";

else if site=**12** then Strata="Отсутствие сахарного диабета, женщина, младше 35 лет, не курит";

else if site=**13** then Strata="Наличие сахарного диабета, женщина, старше 35 лет, не курит";

else if site=**14** then Strata="Отсутствие сахарного диабета, мужчина, старше 35 лет, не курит";

else if site=**15** then Strata="Наличие сахарного диабета, мужчина, старше 35 лет, не курит";

else if site=**16** then Strata="Отсутствие сахарного диабета, женщина, старше 35 лет, не курит";

run; /\*Делаем список более наглядным с названиями страт\*/

%end;

**%mend** rnd;

%***rnd***(seed=**52753033**); /\*Запускаем построение, указывая некоторую случайную величину\*/

**Приложение 2**

option mprint mlogic;

%let amount\_strata=16; /\*Количество страт\*/

%let amount\_group=2; /\*Количество групп терапии\*/

%let amount\_patients=64; /\*Количество участников\*/

%let weight=0.0625, 0.0625, 0.0625, 0.0625, 0.0625, 0.0625, 0.0625, 0.0625, 0.0625, 0.0625, 0.0625, 0.0625, 0.0625, 0.0625, 0.0625, 0.0625; /\*Ожидаемое распределение по стратам\*/

ods results;

/\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*/

/\*\*\*\* Заполнение рандомизационных списков \*\*\*\*\*/

/\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*/

**%macro** ***rnd\_list***;

%do k=**1** %to **150**; /\*Количество заполняемых ранд списков\*/

%do i=**1** %to **1000**; /\*Количество заполнений для каждого рандомизационного списка\*/

data rannum (keep=site);

rannum=ranuni(round(RAND('NORMAL',**10**,**100**),**1.**));

seed=int(rannum\***1000000**);

%do j=**1** %to &amount\_patients;

site=rantbl(seed, &weight);

j=&j;

output;

%end;

run; /\*Присвоение страты каждому из пришедших участников, в зависимости от веса\*/

proc freq data=rannum noprint;

table site/ out = Freq;

run; /\*Подсчет участников в каждой из страт\*/

data rand;

set lib.rnd\_&k;

by site notsorted;

retain num;

if first.site then num=**1**;

else num=num+**1**;

run; /\*Нумерация каждого из созданных списков по пациентов в каждой из страт\*/

data rand;

merge rand Freq;

by site;

if num>COUNT then delete;

run; /\*Присвоение каждой из страт определенного выше числа пациентов\*/

proc freq data=rand noprint;

table med\_grp/ out = GRFreq;

run; /\*Подсчет участников в каждой группе терапии в каждой страте\*/

%if &i=**1** %then %do;

data GRFreqAll\_&k;

set GRFreq;

freq=COUNT-&amount\_patients/**3**;

i=&i;

run; /\*Подсчет отклонений для каждой из групп терапии в каждом из созданных распределений для первого списка\*/

%end;

%if &i>**1** %then %do;

data GRFreq;

set GRFreq;

freq=COUNT-&amount\_patients/**3**;

i=&i;

run; /\*Подсчет отклонений для каждой из групп терапии в каждом из созданных распределений для последующих списков\*/

data GRFreqAll\_&k;

merge GRFreqAll\_&k GRFreq;

by i;

run; /\*Склеивание всех сделанных возможных исходов дисбаланса\*/

%end;

%end;

proc sort data=GRFreqAll\_&k;

by i med\_grp;

run; /\*Сортировка для правильного транспонирования\*/

proc transpose data=GRFreqAll\_&k(drop=percent freq) out=All\_&k;

id med\_grp;

by i;

run; /\*Транспонирование датасета для подсчета дисбаланса\*/

data all\_&k;

set all\_&k;

k=k;

coef=Group\_1/Group\_2;

Disb=Group\_1-Group\_2;

run; /\*Подсчет дисбаланса и коэффициентов распределения\*/

proc ttest data = all\_&k h0 = **1**;

var coef;

run; /\*Проверка каждого из списков по отдельности с помощью t-теста\*/

proc means data = all\_&k noprint;

var Disb;

output out=means\_&k n = coln\_N mean = coln\_mean std=coln\_StdDev median = coln\_Median min=coln\_Min max=coln\_Max lclm=coln\_lclm uclm=coln\_uclm;

run; /\*Подсчет описательной статистики для дисбаланса каждого из списков по отдельности \*/

%if &k=**1** %then %do;

data All;

set all\_&k;

run;

%end;

%else %if &k>**1** %then %do;

data All;

merge All all\_&k;

by k;

run;

%end; /\*Склеивание всех списков воедино\*/

%end;

proc ttest data = all h0 = **1**;

var coef;

run; /\*Проверка всех вместе списков с помощью t-теста\*/

proc means data = all noprint;

var Disb ;

output out=means\_&k n = coln\_N mean = coln\_mean std=coln\_StdDev median = coln\_Median min=coln\_Min max=coln\_Max lclm=coln\_lclm uclm=coln\_uclm;

run; /\*Подсчет описательной статистики для дисбаланса\*/

**%mend** rnd\_list;

%***rnd\_list***; /\*Запуск генерации распределений\*/