Расчет объема выборки

Касьянова Мария

2024-12-09

# 1. ХОБЛ

Исследование не меньшей эффективности препарата T в сравнении с препаратом R в лечении хронической обструктивной болезни лёгких (ХОБЛ). Первичная конечная точка – абсолютное увеличение индекса Тиффно (ОФВ1/ЖЕЛ) за 3 месяца терапии в сравнении со значениями на 1 визите. Из литературных данных известно, что для препарата R абсолютный прирост индекса за 3 месяца в среднем (M±SD) составлял 7.2 ± 3.1%. Предлагаемая граница не меньшей эффективности – 2%.

Статистическая гипотеза не меньшей эффективности в исследовании была сформулирована следующим образом:

* H0: T – R ≤ δ
* HA: T – R > δ

Где T – среднее увеличение индекса Тиффно за 3 месяца в исследуемой группе, R – частота достижения ремиссии в контрольной группе. δ – граница не меньшей эффективности (-2%).

При расчёте необходимого размера выборки были сделаны следующие допущения:

1. Распределение в группы будет равномерным в соотношении 1:1;
2. Односторонняя величина ошибки I рода (α) составляет 0.025;
3. Мощность исследования должна составлять 80%, соответственно, предельная величина ошибки II рода (β) = 0.2;
4. На основании литературных данных мы ожидаем величину эффекта в группе контрольного препарата 7.2 ± 3.1% (R) и в группие исследуемого препарата также 7.2 ± 3.1% (T).
5. Граница не меньшей эффективности препарата по сравнению с активным контролем равна -2%.

Расчёт размера выборки был проведён в RStudio (версия 2024.09.1 Build 394) c использованием языка R (версия 4.4.2) с помощью функции epi.ssninfc пакета epiR (версия 2.0.77).

epi.ssninfc(  
 treat = 7.2,  
 control = 7.2,  
 sigma = 3.1,  
 delta = 2,  
 n = NA,  
 power = 0.8,  
 alpha = 0.025,  
 r = 1   
)

## $n.total  
## [1] 76  
##   
## $n.treat  
## [1] 38  
##   
## $n.control  
## [1] 38  
##   
## $delta  
## [1] 2  
##   
## $power  
## [1] 0.8

Таким образом, в каждую группу необходимо включить не менее 38 обследуемых, всего – 76 пациентов. С учётом частоты выбывания пациентов в ходе исследования, равной 20%, в каждую группу необходимо набрать не менее 48 человек. В скрининг (с учётом 10% выбывания) необходимо включить не менее 54 обследуемых в каждую группу.

# 2. Рассеянный склероз

Исследование превосходства препарата T над референтным лечением R в лечении рассеянного склероза. Первичный показатель эффективности – доля пациентов, у которых к визиту 3 (12 мес.) нет новых очагов по МРТ, развившихся в период наблюдения. Согласно литературным данным, на фоне применения препарата R доля пациентов без новых очагов составляет 66%. Ожидается, что препарат T будет эффективнее препарата R примерно на 15%, при этом спонсор хотел бы, чтобы граница превосходства была бы равна 1%. Спонсор просит, чтобы в группу T было рандомизировано в 2 раза больше пациентов, чем в группу R.

Статистическая гипотеза превосходства в исследовании была сформулирована следующим образом:

* H0: PT – PR ≤ δ
* HA: PT – PR > δ

Где PT – доля пациентов без новых очагов через 3 месяца в исследуемой группе, PR – доля пациентов без новых очагов через 3 месяца в контрольной группе. δ – граница превосходства (1%).

При расчёте необходимого размера выборки были сделаны следующие допущения:

1. Распределение в группы будет проводиться в соотношении 2:1 (T:R);
2. Односторонняя величина ошибки I рода (α) составляет 0.025;
3. Мощность исследования должна составлять 80%, соответственно, предельная величина ошибки II рода (β) = 0.2;
4. На основании литературных данных мы ожидаем величину эффекта в группе исследуемого препарата ~ 81% (PT) и в группе активного наблюдения на уровне ~ 66% (PC).
5. Граница превосходства препарата по сравнению с активным контролем равна 1%.

Расчёт размера выборки был проведён в RStudio (версия 2024.09.1 Build 394) c использованием языка R (версия 4.4.2) с помощью функции epi.sssupb пакета epiR (версия 2.0.77).

epi.sssupb(  
 treat = 0.81,  
 control = 0.66,  
 delta = 0.01,  
 n = NA,  
 power = 0.8,  
 alpha = 0.025,  
 r = 2  
)

## $n.total  
## [1] 363  
##   
## $n.treat  
## [1] 242  
##   
## $n.control  
## [1] 121  
##   
## $delta  
## [1] 0.01  
##   
## $power  
## [1] 0.8

Таким образом, в исследование необходимо включить не менее 363 обследуемых (242 и 121). С учётом частоты выбывания пациентов в ходе исследования, равной 20%, в группы необходимо набрать не менее 303 и 152 человек. В скрининг (с учётом 10% выбывания) необходимо включить суммарно не менее 506 обследуемых.

# 3. Интраоперационные кровотечения

Исследование превосходства нового препарата А над препаратом Б в профилактике интраоперационных кровотечений. Первичная конечная точка – объём кровопотери в мл. Из данных литературы известно, что средний объём кровопотери при применении препарата Б составляет 306 ± 52 мл. Ожидается, что препарат А в среднем будет снижать объём кровопотери на 20% в сравнении с препаратом Б. Границу превосходства предполагается установить равной 50 мл. Спонсор просит рассчитать объём выборок таким образом, чтобы препарат А получили 75% пациентов, включённых в исследование.

Статистическая гипотеза превосходства в исследовании была сформулирована следующим образом:

* H0: mA – mB ⩾ δ
* HA: mA – mB < δ

Где mA – средний объём кровопотери при применении препарата А, mB – средний объём кровопотери при применении препарата Б, δ – граница превосходства (-50).

При расчёте необходимого размера выборки были сделаны следующие допущения:

1. Распределение в группы будет проводиться в соотношении 3:1 (А:Б);
2. Односторонняя величина ошибки I рода (α) составляет 0.025;
3. Мощность исследования должна составлять 80%, соответственно, предельная величина ошибки II рода (β) = 0.2;
4. На основании литературных данных мы ожидаем величину эффекта в группе исследуемого препарата ~ 244.8 ± 52 мл (mА) и в группе активного наблюдения на уровне ~ 306 ± 52 мл (mB).
5. Граница превосходства препарата по сравнению с активным контролем равна -50

Расчёт размера выборки был проведён в RStudio (версия 2024.09.1 Build 394) c использованием языка R (версия 4.4.2) с помощью функции epi.sssupc пакета epiR (версия 2.0.77).

Так как данная функция не работает с отрицательными значениями границы превосходства, обозначение контрольной и опытной группы были поменяны местами, а нулевая и альтернативные гипотезы были переписана следущим образом:

* H0: mB – mA ≤ -δ
* HA: mB – mA > -δ

epi.sssupc(  
 treat = 306,  
 control = 244.8,  
 sigma = 52,  
 delta = 50,  
 n = NA,  
 power = 0.8,  
 alpha = 0.025,  
 r = 1/3   
)

## $n.total  
## [1] 903  
##   
## $n.treat  
## [1] 226  
##   
## $n.control  
## [1] 677  
##   
## $delta  
## [1] 50  
##   
## $power  
## [1] 0.8

Таким образом, в группу для препарата А необходимо включить не менее 677 обследуемых, а в группу для препарата Б - 226, то есть всего – 903 человека. С учётом частоты выбывания пациентов в ходе исследования, равной 20%, необходимо набрать не менее 847 и 283 человек в две группы соответственно. А в скрининг (с учётом 10% выбывания) необходимо включить всего не менее 1256 обследуемых.

# 4. Псориатический артрит

Известно, что на фоне лечения препаратом Х рецидив псориатического артрита за 6 месяцев наблюдается в 32% случаев. Спонсор хочет исследовать новый препарат У, который, как ожидается, снизит частоту рецидивов на 15%. Границей превосходства можно считать снижение доли пациентов с рецидивом более, чем на 5%. При этом существует сложность. У спонсора в наличии имеется только 250 пачек препарата Х и он просит, чтобы рандомизационное соотношение было таким, чтобы препарата хватило на проведение исследования.

Статистическая гипотеза превосходства в исследовании была сформулирована следующим образом:

* H0: PY – PX ⩾ δ
* HA: PY – PX < δ

Где PY – частота рецидивов в группе препарата Y, PX – частота рецидивов в контрольной группе Х, δ – граница превосходства (-5%).

При расчёте необходимого размера выборки были сделаны следующие допущения:

1. В группе, принимающей препарат Y будет не больше 200 человек (чтобы с учетом 20% выбывания в нее набрали не больше 250 человек);
2. Односторонняя величина ошибки I рода (α) составляет 0.025;
3. Мощность исследования должна составлять 80%, соответственно, предельная величина ошибки II рода (β) = 0.2;
4. На основании литературных данных мы ожидаем величину эффекта в группе исследуемого препарата ~ 17% (PY) и в группе активного наблюдения на уровне ~ 32% (PX).
5. Граница превосходства препарата по сравнению с активным контролем равна -5%.

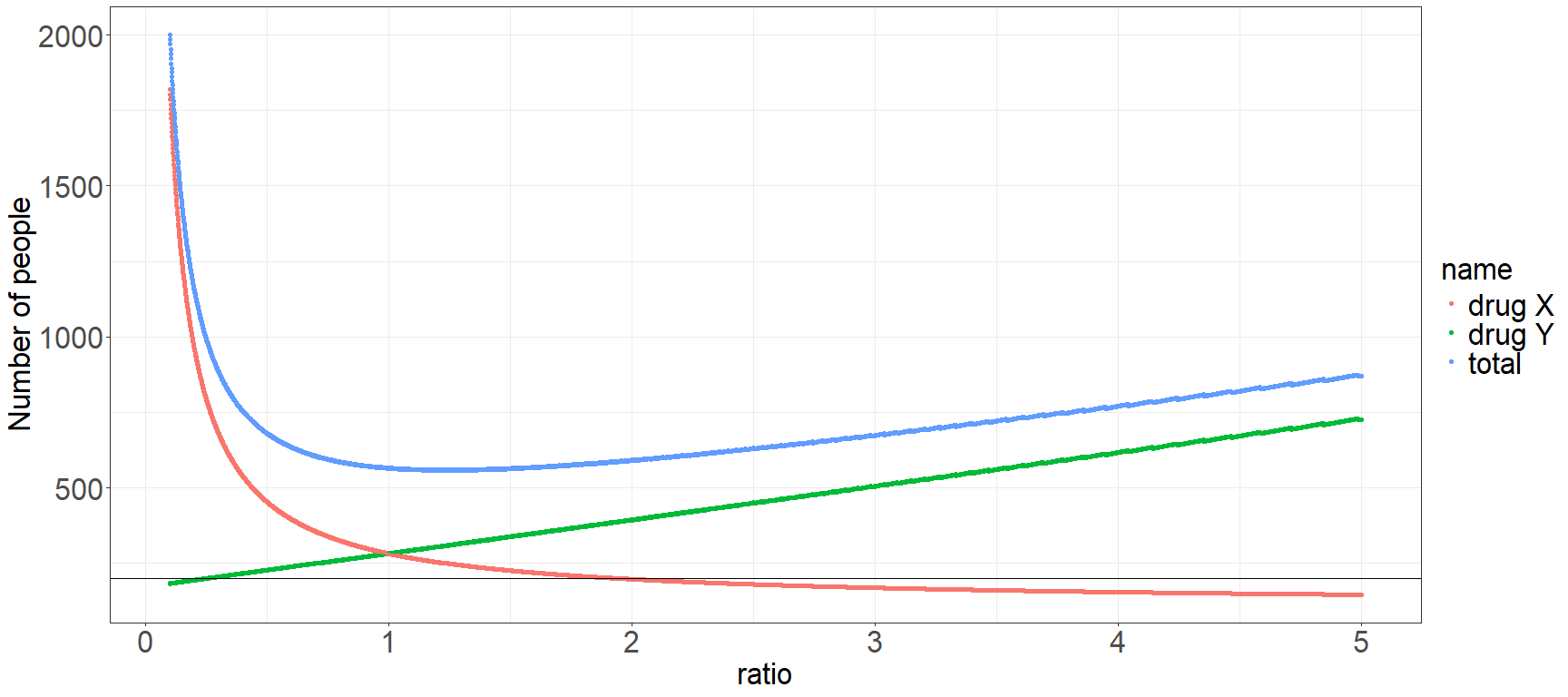
Расчёт размера выборки был проведён в RStudio (версия 2024.09.1 Build 394) c использованием языка R (версия 4.4.2) с помощью функции epi.sssupb пакета epiR (версия 2.0.77) и тестированием нескольких значений параметра r.

Так как данная функция не работает с отрицательными значениями границы превосходства, обозначение контрольной и опытной группы были поменяны местами, а нулевая и альтернативные гипотезы были переписана следущим образом:

* H0: - PY + PX ≤ -δ
* HA: - PY + PX > -δ

сначала изучим, как меняются размеры групп в зависимости от их соотношения:

ratio <- seq(from = 0.1, to = 5, by = 0.001)  
treat <- c()  
control <- c()  
total <- c()  
  
  
for (i in ratio){  
numbers <- epi.sssupb(  
 treat = 0.32,  
 control = 0.17,  
 delta = 0.05,  
 n = NA,  
 power = 0.8,  
 alpha = 0.025,  
 r = i  
)  
treat <- c(treat, numbers$n.treat)  
control <- c(control, numbers$n.control)  
total <- c(total, numbers$n.total)  
}  
  
  
tibble(ratio = ratio, `drug Y` = treat, `drug X` = control, total = total)%>%  
 pivot\_longer(c(`drug Y`, `drug X`, total), values\_to ="Number of people")%>%  
 ggplot(aes(x = ratio, y = `Number of people`, color = name))+  
 geom\_point()+  
 theme\_bw()+  
 geom\_hline(yintercept = 200)+  
 theme(  
 axis.text = element\_text(size = 25),  
 axis.title = element\_text(size = 25),  
 legend.title = element\_text(size = 25),  
 legend.text = element\_text(size = 25)  
 )



Видно, что при увеличении параметра r, уменьшается размер группы, принимающей препарат Х (порог в 200 человек отмечен черной линией), но начиная с некоторого момента возрастает общее количество исследуемых. Соответвенно те значения r, которые не вошли в график, нам не подойдут, и можно выбрать наилучший из рассмотренных вариантов:

tibble(ratio = ratio, `drug Y` = treat, `drug X` = control, total = total)%>%  
 filter(`drug X` <= 200)%>%  
 arrange(total)%>%  
 slice(1)

## # A tibble: 1 × 4  
## ratio `drug Y` `drug X` total  
## <dbl> <dbl> <dbl> <dbl>  
## 1 1.91 383 200 583

Таким образом, в группы необходимо включить не менее 200 и 383 обследуемых, всего – 583 пациента. С учётом частоты выбывания пациентов в ходе исследования, равной 20%, в первую группу нужно отобрать 250 человек, а во вторую 479. В скрининг (с учётом 10% выбывания) необходимо включить не менее 810 людей суммарно.