eWHORM – Statistical considerations for setting up the simulation study

MUW Team

Table of contents

Preamble

Brief report to summarise the key points of the planned design and discuss the simulation study.

Aims and scope

Preliminary study on the design of a randomized, placebo-controlled, double-blind, parallel-group phase-II study to investigate the efficacy and safety of oxfendazole (OXF) in subjects with onchocerciasis, loiasis, mansonellosis and/or soil-transmitted helminth (STH).

In particular, the goal of this report is to explore the state-of-the-art and establish the basis for a simulation study aiming at investigating the key operating characteristics of the trial.

The simulations will optimize different design elements, including interim analyses and stopping rules, dose selection at interim, (adaptive) randomization and treatment allocation, sample size reassessment and subgroup analyses. Furthermore, the simulations will help in guaranteeing a prespecified statistical power, type 1 error rate control and ensuring the trial integrity and validity. Adaptive decision rules will be pre-specified and interim and final analysis will feed into the design. With specific statistical testing and estimation procedures accounting for adaptations such as sample size reassessment and dropping of treatment arms, the type 1 error rate will be controlled and the sample size reassessed.

Main elements of the design

Sub-studies

Focusing on filarial and STH infections, we consider a basket trial to test OXF against four diseases (onchocerciasis, loiasis, mansonellosis and STH). The basket trial then considers four sub-studies corresponding to each of the previous-mentioned diseases.

Endpoints

The efficacy endpoints of the sub-studies are the following:

- OXF in Onchocerciasis patients
  - The efficacy endpoint will be the presence (or absence) of live female adult worms with normal embryogenesis at month 12.
- OXF in Loiasis patients:
  - The efficacy endpoint will be the presence (or presence) of L. loa mf at 12 months.

- OXF in mansonellosis patients:
  - The efficacy endpoint will be the presence (or absence) of mf at month 12.
- OXF in T. trichiura patients:
  - The efficacy endpoint is the T. trichiura cure rate assessed at 14 - 21 days post-treatment using the duplicate Kato-Katz technique. Other endpoints are egg reduction rates (ERRs) and infection status and intensity of infection.

Notes: 1) For some of the diseases it will be important to perform subgroup analyses (for now not included here). 2) For mansonellosis, we will need to account for co-infections. In particular, mansonellosis impacts co-infections with M. tuberculosis and HIV.

Treatments

In the three sub-studies with filarial patients (onchocerciasis, loiasis and mansonellosis), patients will receive either 400 mg OXF single dose (SD), 800 mg OXF SD, 800 mg OXF once daily for 5 days, or placebo. In the sub-study with trichuriasis patients, OXF will be administered at 400mg and 800mg SD.

- onchocerciasis, loiasis and mansonellosis: three treatments and placebo
- trichuriasis patients: two treatments and placebo

Trial stages

The trial includes two-stages.

- Stage 1. In the first stage, 120 patients per indication will be included. After stage 1 ends, an interim analysis will be performed in all sub-studies. According to that, the trial may be adapted (treatment arms can be stopped for futility or efficacy, sample size will be reassessed).
- Stage 2. Trial continues to the second stage where up to 80 additional patients may be allocated to the selected doses and the placebo group.

Study duration will be up to 12 months, depending on the indication.

Trial objectives

The main objectives of the trial are the following:

- Primary objective: To demonstrate a positive dose-response relationship of OXF in the primary endpoint for each of the four indications: onchocerciasis, loiasis, mansonellosis, trichuriasis/STH infection.

- Secondary objectives: To demonstrate the efficacy of specific doses.

Further points to be considered include (not all included for now): the safety and tolerability of OXF in helminth infected individuals, geographical region, PK parameters of OXF and their relationship to efficacy response and safety and tolerability will be investigated.

Design considerations in previous and/or published studies

Summary clinical papers

- Welsche et al. (2022)
  - The primary outcome was egg reduction rate (ERR) of T trichiura 14–21 days after treatment in the available case population.
  - The estimated sample size, based on simulations, resulted in 160 participants would be needed per group to ensure a power of 90% that the upper limit of the two-sided 95% CIs would exclude a difference of more than 2 percentage points (the non-inferiority margin) in favour of the ivermectin and albendazole combination group, assuming a true ERR of 98% in both groups.
  - Analysis:
    * ERRs were calculated using geometric mean egg counts
    * To estimate the difference between ERRs and 95% CIs, a bootstrap resampling method with 5000 replicates was used.
    * Logistic regression with adjustment for baseline infection intensity, age, sex, and weight was also performed. For each follow-up timepoint, ERRs and cure rates were calculated.
- Hürlimann, et al. (2022)
  - The primary outcome was cure rate against T trichiura, defined as the proportion of participants with no eggs in their faeces 14–21 days after treatment, assessed by Kato-Katz thick smears
  - Secondary outcomes were the ERR against T trichiura; cure rates and ERRs against A lumbricoides, hookworm, and S stercoralis; and infection status assessed by qPCR
  - Sample size based on the unpooled difference in proportions sample size formula. 143 participants per treatment group
  - Analysis:
    * A melded binomial test with mid-p correction was used to calculate differences in cure rates between the two treatment groups and to estimate the corresponding confidence intervals.
    * Geometric mean and arithmetic mean egg counts, and bootstrap as before.
- Gardon, et al. (2002)

  - The primary outcome was the vital status of female worms as assessed from the proportion and mean number of female worms in the MD category per nodule.
  - For an expected difference of 30% in the proportion of moribund and dead female worms between the four groups at the end of the trial, and for a patient drop-out rate of 50%, 544 patients were needed, taking into account the data in the published work for the usual proportion of dead worms found before treatment
  - Analysis:
    * calculated the mean numbers of microfilariae per mg of skin with William's geometric mean
    * tested for any difference between all the four groups, with the chi2 test for proportions, and with ANOVA or the Kruskal-Wallis rank-sum test for means. Random-effects logistic-regression models to compare the effects of the three other treatment regimens with that of the control group.

- Preprint paper pending to be included:
  - Zoleko-Manego et al on loiasis in adult patients in Gabon
- Trial protocol pending to be included

Preliminary code to simulate data

Next, we include a preliminary function to simulate basket trial data. In particular, the generated individual data includes the binary response and as covariates

- Substudy: we considered four substudies, which would correspond to each of the diseases in the basket trial
- Treatment: we considered two treatment arms and a control arm
- Patient's covariates: we included sex as covariate

The covariates treatment and sex are considered as fixed effects, while substudy is considered as random effect in this example.

```
simulate_binary_data <- function(n, beta, b, ntreat=3, nsex=2, nstudies=4) {
  # add intercept and design matrix for fixed effects
  x <- matrix(rnorm(n), nrow = n, ncol = 1)
  if (ntreat==3) {
    treatments <- factor(sample(1:ntreat, n, replace = TRUE),
                         levels = 1:ntreat,
                         labels = c("Placebo", "Treatment 1", "Treatment 2"))
    X <- model.matrix(~ treatments - 1)
    x <- cbind(x, X)
  }
  if (nsex==2) {
    sex <- factor(sample(c("Male", "Female"), n, replace = TRUE))
    X <- model.matrix(~ sex - 1)
    x <- cbind(x, X)
  }
  # add random effects
  if (nstudies==4) {
    studies <- factor(sample(1:nstudies, n, replace = TRUE),
                      levels = 1:nstudies,
                      labels = c("S1", "S2", "S3", "S4"))
    Z <- model.matrix(~ studies - 1)
    rn = rnorm(n, mean = 0, sd = 1)
    z <-  Z*rn
  }

  # compute linear predictor as a function of covariates and coefficients
  eta <- x %*% beta + z %*% b

  # convert linear predictor to probability using logistic function
```

```
  p <- plogis(eta)

  # simulate binary outcomes based on probability
  y <- rbinom(n, 1, p)

  # simulated data
  data <- data.frame(y, x, Z)

  if (ntreat==3 && nsex==2 && nstudies==4) {
    data$treatments <- "T"
    data$treatments[data$treatmentsPlacebo==1] <- "Control"
    data$treatments[data$treatmentsTreatment.1==1] <- "Treat1"
    data$treatments[data$treatmentsTreatment.2==1] <- "Treat2"
    data$treatmentsPlacebo<-NULL
    data$treatmentsTreatment.1<-NULL
    data$treatmentsTreatment.2<-NULL
    data$sex <- "S"
    data$sex[data$sexFemale==1] <- "F"
    data$sex[data$sexMale==1] <- "M"
    data$sexFemale<-NULL
    data$sexMale<-NULL
    data$substudy <- "SS"
    data$substudy[data$studiesS1==1] <- "Study1"
    data$substudy[data$studiesS2==1] <- "Study2"
    data$substudy[data$studiesS3==1] <- "Study3"
    data$substudy[data$studiesS4==1] <- "Study4"
    data$studiesS1<-NULL
    data$studiesS2<-NULL
    data$studiesS3<-NULL
    data$studiesS4<-NULL
    data$V1<-NULL
  }

  model <- glmer(y ~ treatments + sex + (1 | substudy),
                 data = data, family = binomial)

  return(list(data=data, model=model))

}
```

Example:

```r
set.seed(2345)
ntreat=3; nsex=2; nstudies=4; n =4*1000
beta <- c(1,0,0.5,1,0,0.2)
b=c(.1,.2,.4,.5)
#
sim_res <- simulate_binary_data(n=n, beta=beta, b=b, ntreat=3, nsex=2, nstudies=4)
head(sim_res$data,10)
```

```
   y treatments sex substudy
1  1     Treat2   F   Study4
2  1    Control   M   Study3
3  1     Treat2   F   Study1
4  0     Treat1   F   Study2
5  1     Treat1   M   Study3
6  0    Control   M   Study1
7  0     Treat2   M   Study3
8  0     Treat2   M   Study2
9  1     Treat2   M   Study2
10 1     Treat1   M   Study1
```

Fitting the model

```r
summary(sim_res$model)
```

```
Generalized linear mixed model fit by maximum likelihood (Laplace
  Approximation) [glmerMod]
 Family: binomial  ( logit )
Formula: y ~ treatments + sex + (1 | substudy)
   Data: data

     AIC      BIC   logLik deviance df.resid
  5251.9   5283.3  -2620.9   5241.9     3995

Scaled residuals:
    Min      1Q  Median      3Q     Max
-1.6298 -1.0877  0.6166  0.8427  1.0160

Random effects:
 Groups   Name        Variance  Std.Dev.
 substudy (Intercept) 0.0002814 0.01677
```

```
Number of obs: 4000, groups:  substudy, 4

Fixed effects:
                 Estimate Std. Error z value Pr(>|z|)
(Intercept)      -0.02474    0.06423  -0.385  0.70005
treatmentsTreat1  0.36416    0.07849   4.639 3.49e-06 ***
treatmentsTreat2  0.80871    0.08134   9.942  < 2e-16 ***
sexM              0.19008    0.06578   2.890  0.00386 **
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Correlation of Fixed Effects:
            (Intr) trtmT1 trtmT2
trtmntsTrt1 -0.593
trtmntsTrt2 -0.583  0.468
sexM        -0.509  0.002  0.024
```

```
ranef(sim_res$model)
```

```
$substudy
        (Intercept)
Study1 -0.007099705
Study2  0.002892139
Study3  0.001313538
Study4  0.002887297

with conditional variances for "substudy"
```