Mediana Designer's project files

Mediana Designer Community Edition is a free software tool that supports traditional and simulation-based power/sample size calculations in fixed-sample and group-sequential trials. To learn more about Mediana Designer Community Edition, visit the Biopharmnet web site at

http://biopharmnet.com/mediana-designer/

Project files

Multiple project files have been created to demonstrate how Mediana Designer can be used to perform power and sample size calculations in late-stage clinical trials. The project files can be downloaded from

http://biopharmnet.com/mediana-designer-project-files/

Detailed descriptions of the project files/case studies are provided below:

- TraditionalSampleSize module: 14 project files have been set up to show how to perform analytical (frequentist) and simulation-based (Bayesian) power calculations in fixed-design trials.
- TraditionalSimulations module: 16 project files have been set up to show how to perform simulation-based power calculations in fixed-sample trials using the Clinical Scenario Evaluation.
- GroupSequential module: 20 project files have been set up to show how to perform analytical and simulation-based power calculations in group-sequential trials.

For more information on the statistical methods implemented in the individual modules, see the Mediana Designer Technical Manual (http://biopharmnet.com/software/mediana-designer-manual.pdf).

TraditionalSampleSize module

Case study 1

This case study illustrates analytical power calculations in two-arm clinical trials with a superiority objective when the primary endpoint is normally distributed. It is based on a Phase III clinical trial for the treatment of schizophrenia. A single dose of the experimental treatment is tested versus placebo using a 1:1 randomization scheme. The primary endpoint is the change in the Positive and Negative Syndrome Scale (PANSS) total score and a reduction in the total score is associated with improvement. This endpoint follows a normal distribution and the primary analysis is performed using the two-sample Z-test to show that the treatment is superior to placebo. The goal is to compute power of the primary test when the total sample size in the trial is 440 patients.

Case study 2

This case study illustrates analytical power calculations and simulation-based assurance calculations in two-arm clinical trials with a superiority objective when the primary endpoint is normally distributed. It is based on a Phase III clinical trial for the treatment of schizophrenia. A single dose of the experimental treatment is tested versus placebo using a 1:1 randomization scheme. The primary endpoint is the change in the Positive and Negative Syndrome Scale (PANSS) total score and a reduction in the total score is associated with improvement. This endpoint follows a normal distribution and the primary analysis is performed using the two-sample Z-test to show that the treatment is superior to placebo. The goal is to compute power of the primary test when the total sample size in the trial is 440 patients. In addition, the assurance calculations are performed with the total sample size of 440 patients assuming vague priors for the means and standard deviations in the placebo and treatment arms. These priors are derived from a relatively small historical trial with 50 patients in each arm.

Case study 3

This case study illustrates analytical sample size calculations and simulation-based assurance calculations in two-arm clinical trials with a superiority objective when the primary endpoint is normally distributed. It is based on a Phase III clinical trial for the treatment of schizophrenia. A single dose of the experimental treatment is tested versus placebo using a 1:1 randomization scheme. The primary endpoint is the change in the Positive and Negative Syndrome Scale (PANSS) total score and a reduction in the total score is associated with improvement. This endpoint follows a normal distribution and the primary analysis is performed using the two-sample Z-test to show that the treatment is superior to placebo. The goal is to compute the total sample size corresponding to 90% power. The assurance calculations are performed using the resulting sample size and assuming vague priors for the means and standard deviations in the placebo and treatment arms. These priors are derived from a relatively small historical trial with 50 patients in each arm.

Case study 4

This case study illustrates analytical power calculations and simulation-based assurance in two-arm clinical trials with a non-inferiority objective when the primary endpoint is normally distributed. It is based on a Phase III clinical trial for the treatment of schizophrenia. A single dose of the experimental treatment is tested versus an active control using a 1:1 randomization scheme. The primary endpoint is the change in the Positive and Negative Syndrome Scale (PANSS) total score and a reduction in the total score is associated with improvement. This endpoint follows a normal distribution and the primary analysis is performed using the two-sample Z-test to show that the treatment is non-inferior to the control. The non-inferiority margin is 4 (it is important to note that the non-inferiority margin is positive when a lower value of the primary endpoint indicates a beneficial effect). The goal is to compute power of the non-inferiority test when the total sample size in the trial is 440 patients. In addition, the assurance calculations are performed using the resulting sample size and assuming vague priors for the means and standard deviations in the placebo and treatment arms. These priors are derived from a relatively small historical trial with 50 patients in each arm.

Case study 5

This case study illustrates analytical sample size calculations in two-arm clinical trials with a superiority objective when the primary endpoint is binary. It is based on a Phase III clinical trial in patients with rheumatoid arthritis. There are two arms in the trial (experimental treatment versus placebo) and an unbalanced design with a 2:1 randomization scheme (treatment versus placebo) is employed. The primary endpoint is a binary outcome (response versus no response) based on the American College of Rheumatology (ACR50) definition of improvement. A higher response rate indicates a beneficial treatment effect. The primary analysis is performed using the two-sample Z-test for proportions to show that the treatment is superior to placebo. The goal is to compute the total sample size corresponding to 90% power.

Case study 6

This case study illustrates analytical sample size calculations and simulation-based assurance calculations in two-arm clinical trials with a superiority objective when the primary endpoint is binary. It is based on a Phase III clinical trial in patients with rheumatoid arthritis. There are two arms in the trial (experimental treatment versus placebo) and an unbalanced design with a 2:1 randomization scheme (treatment versus placebo) is employed. The primary endpoint is a binary outcome (response versus no response) based on the American College of Rheumatology (ACR50) definition of improvement. A higher response rate indicates a beneficial treatment effect. The primary analysis is performed using the two-sample Z-test for proportions to show that the treatment is superior to placebo. The goal is to compute the total sample size corresponding to 90% power. The assurance calculations are performed using the resulting sample size and the prior distributions for the ACR50 response rates in the control and treatment arms that are derived from a Phase II trial with 100 patients in each arm.

Case study 7

This case study illustrates analytical sample size calculations and simulation-based assurance calculations in two-arm clinical trials with a non-inferiority objective when the primary endpoint is binary. It is based on a Phase III clinical trial in patients with rheumatoid arthritis. There are two arms in the trial (experimental treatment versus active control) and an unbalanced design with a 2:1 randomization scheme (treatment versus control) is employed. The primary endpoint is a binary outcome (response versus no response) based on the American College of Rheumatology (ACR50) definition of improvement. A higher response rate indicates a beneficial treatment effect. The primary analysis is performed using the two-sample Z-test for proportions to show that the treatment is non-inferior to the active control. The non-inferiority margin is -20% or -0.2 (note that the non-inferiority margin is negative if a higher value of the primary endpoint indicates a beneficial effect). The goal is to compute the total sample size corresponding to 90% power. The assurance calculations are performed using the resulting sample size and the prior distributions for the ACR50 response rates in the control and treatment arms that are derived from a Phase II trial with 100 patients in each arm.

Case study 8

This case study illustrates analytical power calculations in two-arm clinical trials with a superiority objective when the primary endpoint is a time-to-event endpoint. It is based on a Phase III clinical trial in patients with metastatic colorectal cancer. An unbalanced design with 2:1 randomization (experimental treatment versus best supportive care) is utilized in the trial. The primary endpoint is the time to disease progression or death and is assumed to follow an exponential distribution. The primary analysis is performed using the log-rank test to show that the treatment is superior to the control. The goal is to compute power corresponding to 200 events without accounting for loss to follow-up or administrative censoring (in which case the number of events is equal to the sample size).

Case study 9

This case study illustrates analytical power calculations and simulation-based assurance calculations in two-arm clinical trials with a superiority objective when the primary endpoint is a time-to-event endpoint. It is based on a Phase III clinical trial in patients with metastatic colorectal cancer. An unbalanced design with 2:1 randomization (experimental treatment versus best supportive care) is utilized in the trial. The primary endpoint is the time to disease progression or death and is assumed to follow an exponential distribution. The primary analysis is performed using the log-rank test to show that the treatment is superior to the control. The goal is to compute power corresponding to 200 events without accounting for loss to follow-up or administrative censoring (in which case the number of events is equal to the sample size). The assurance calculations are performed using the prior distributions for the hazard rates in the control and treatment arms that are derived from a historical trial with 50 events in the control arm and 40 events in the treatment arm.

Case study 10

This case study illustrates analytical event count calculations and simulation-based assurance calculations in two-arm clinical trials with a superiority objective when the primary endpoint is a time-to-event endpoint. It is based on a Phase III clinical trial in patients with metastatic colorectal cancer. An unbalanced design with 2:1 randomization (experimental treatment versus best supportive care) is utilized in the trial. The primary endpoint is the time to disease progression or death and is assumed to follow an exponential distribution. The primary analysis is performed using the log-rank test to show that the treatment is superior to the control. The goal is to compute the required number of events without accounting for loss to follow-up or administrative censoring (in which case the number of events is equal to the sample size) to achieve 90% power. The assurance calculations are performed using the prior distributions for the hazard rates in the control and treatment arms that are derived from a historical trial with 50 events in the control arm and 40 events in the treatment arm.

Case study 11

This case study illustrates analytical event count calculations and simulation-based assurance calculations in two-arm clinical trials with a non-inferiority objective when the primary endpoint is a time-to-event endpoint. It is based on a Phase III clinical trial in patients with metastatic colorectal cancer. An unbalanced design with 2:1 randomization (experimental treatment versus active control) is utilized in the trial. The primary endpoint is the time to disease progression or death and is assumed to follow an exponential distribution. The primary analysis is performed using the log-rank test to show that the treatment is non-inferior to the active control. The non-inferiority margin on the hazard ratio scale is 1.3 (note that the non-inferiority margin needs to be greater than 1 since a larger value of the primary endpoint indicates a beneficial effect). The goal is to compute the required number of events without accounting for loss to follow-up or administrative censoring (in which case the number of events is equal to the sample size) to achieve 90% power. The assurance calculations are performed using the prior distributions for the hazard rates in the control and treatment arms that are derived from a historical trial with 50 events in the control arm and 40 events in the treatment arm.

Case study 12

This case study illustrates analytical sample size calculations and simulation-based assurance calculations in two-arm clinical trials with a superiority objective when the primary endpoint is a time-to-event endpoint and the patient accrual/dropout processes are taken into account. It is based on a Phase III clinical trial in patients with metastatic colorectal cancer. An unbalanced design with 2:1 randomization (experimental treatment versus best supportive care) is utilized in the trial. The primary endpoint is the time to disease progression or death and is assumed to follow an exponential distribution. The primary analysis is performed using the log-rank test to show that the treatment is superior to the control. The trial includes a 9-month accrual period (patients are assumed to be recruited at a uniform rate) with a minimum follow-up period of 12 months and the annual dropout rate is 5% (time to dropout is assumed to be exponentially distributed). The goal is to compute the total sample size (number of enrolled patients) to achieve 90% power. In addition, the assurance calculations are performed using the prior distributions for the hazard rates in the control

and treatment arms that are derived from a historical trial with 50 events in the control arm and 40 events in the treatment arm.

Case study 13

This case study illustrates analytical sample size calculations and simulation-based assurance calculations in two-arm clinical trials with a superiority objective when the primary endpoint is a time-to-event endpoint and the patient accrual/dropout processes are taken into account. It is based on a Phase III clinical trial in patients with metastatic colorectal cancer. An unbalanced design with 2:1 randomization (experimental treatment versus best supportive care) is utilized in the trial. The primary endpoint is the time to disease progression or death and is assumed to follow an exponential distribution. The primary analysis is performed using the log-rank test to show that the treatment is superior to the control. The trial includes a 9-month accrual period with a minimum follow-up period of 12 months. The patient accrual is governed by a truncated exponential distribution with the median accrual time of 6 months. The annual dropout rate is 5% (time to dropout is assumed to be exponentially distributed). The goal is to compute the total sample size (number of enrolled patients) to achieve 90% power. In addition, the assurance calculations are performed using the prior distributions for the hazard rates in the control and treatment arms that are derived from a historical trial with 50 events in the control arm and 40 events in the treatment arm.

Case study 14

This case study illustrates analytical sample size calculations and simulation-based assurance calculations in two-arm clinical trials with a non-inferiority objective when the primary endpoint is a time-to-event endpoint and the patient accrual/dropout processes are taken into account. It is based on a Phase III clinical trial in patients with metastatic colorectal cancer. An unbalanced design with 2:1 randomization (experimental treatment versus active control) is utilized in the trial. The primary endpoint is the time to disease progression or death and is assumed to follow an exponential distribution. The primary analysis is performed using the log-rank test to show that the treatment is non-inferior to the active control. The non-inferiority margin on the hazard ratio scale is 1.3 (note that the non-inferiority margin needs to be greater than 1 since a larger value of the primary endpoint indicates a beneficial effect). The trial includes a 9-month accrual period (patients are assumed to be recruited at a uniform rate) with a minimum follow-up period of 12 months and the annual dropout rate is 5% (time to dropout is assumed to be exponentially distributed). The goal is to compute the total sample size (number of enrolled patients) to achieve 90% power. In addition, the assurance calculations are performed using the prior distributions for the hazard rates in the control and treatment arms that are derived from a historical trial with 50 events in the control arm and 40 events in the treatment arm.

TraditionalSimulations module

Case study 1

This case study illustrates simulation-based power calculations in two-arm clinical trials with a superiority objective when the primary endpoint is normally distributed. It is based on the same clinical trial for the treatment of schizophrenia as Case study 1 for the TraditionalSampleSize module. A single dose of the experimental treatment is tested versus placebo using a 1:1 randomization scheme. The primary endpoint is the change in the Positive and Negative Syndrome Scale (PANSS) total score and a reduction in the total score is associated with improvement. This endpoint follows a normal distribution and the primary analysis is performed using the two-sample Z-test to show that the treatment is superior to placebo. The goal is to compute power of the primary test as a function of the total sample size in the trial (total sample size ranges from 400 to 500).

Case study 2

This case study illustrates simulation-based power calculations in two-arm clinical trials with a superiority objective when the primary endpoint is normally distributed. It is based on the same clinical trial for the treatment of schizophrenia as Case study 2 for the TraditionalSampleSize module. A single dose of the experimental treatment is tested versus placebo using a 1:1 randomization scheme. The primary endpoint is the change in the Positive and Negative Syndrome Scale (PANSS) total score and a reduction in the total score is associated with improvement. This endpoint follows a normal distribution and the primary analysis is performed using the two-sample Z-test to show that the treatment is superior to placebo. The goal is to compute assurance as a function of the total sample size in the trial (total sample size ranges from 400 to 500). The assurance calculations are performed based on vague prior distributions for the means and standard deviations in the placebo and treatment arms. These priors are derived from a relatively small historical trial with 50 patients in each arm.

Case study 3

This case study illustrates simulation-based power calculations in two-arm clinical trials with a non-inferiority objective when the primary endpoint is normally distributed. It is based on the same clinical trial for the treatment of schizophrenia as Case study 4 for the TraditionalSampleSize module. A single dose of the experimental treatment is tested versus an active control using a 1:1 randomization scheme. The primary endpoint is the change in the Positive and Negative Syndrome Scale (PANSS) total score and a reduction in the total score is associated with improvement. This endpoint follows a normal distribution and the primary analysis is performed using the two-sample Z-test to show that the treatment is non-inferior to the control. The non-inferiority margin is 4 (non-inferiority margin is positive since a lower value of the primary endpoint indicates a beneficial effect). The goal is to compute power as a function of the total sample size in the trial (total sample size ranges from 400 to 500).

Case study 4

This case study illustrates simulation-based power calculations in two-arm clinical trials with a superiority objective when the primary endpoint is binary. It is based on the same clinical trial for the treatment of rheumatoid arthritis as Case study 5 for the TraditionalSampleSize module. There are two arms in the trial (experimental treatment versus placebo) and an unbalanced design with a 2:1 randomization scheme (treatment versus placebo) is employed. The primary endpoint is a binary outcome based on the American College of Rheumatology (ACR50) definition of improvement. A higher response rate indicates a beneficial treatment effect. The primary analysis is performed using the two-sample Z-test for proportions to show that the treatment is superior to placebo. The goal is to compute power as a function of the total sample size in the trial (total sample size ranges from 210 to 300).

Case study 5

This case study illustrates simulation-based power calculations in two-arm clinical trials with a superiority objective when the primary endpoint is binary. It is based on the same clinical trial for the treatment of rheumatoid arthritis as Case study 6 for the TraditionalSampleSize module. There are two arms in the trial (experimental treatment versus placebo) and an unbalanced design with a 2:1 randomization scheme (treatment versus placebo) is employed. The primary endpoint is a binary outcome based on the American College of Rheumatology (ACR50) definition of improvement. A higher response rate indicates a beneficial treatment effect. The primary analysis is performed using the two-sample Z-test for proportions to show that the treatment is superior to placebo. The goal is to compute assurance as a function of the total sample size in the trial (total sample size ranges from 210 to 300). The assurance calculations are performed using the prior distributions for the ACR50 response rates in the control and treatment arms that are derived from a Phase II trial with 100 patients in each arm.

Case study 6

This case study illustrates simulation-based power calculations in two-arm clinical trials with a non-inferiority objective when the primary endpoint is binary. It is based on the same clinical trial for the treatment of rheumatoid arthritis as Case study 7 for the TraditionalSampleSize module. There are two arms in the trial (experimental treatment versus active control) and an unbalanced design with a 2:1 randomization scheme (treatment versus placebo) is employed. The primary endpoint is a binary outcome based on the American College of Rheumatology (ACR50) definition of improvement. A higher response rate indicates a beneficial treatment effect. The primary analysis is performed using the two-sample Z-test for proportions to show that the treatment is non-inferior to the active control. The non-inferiority margin is set to -20% or -0.2 (non-inferiority margin is negative since a higher value of the primary endpoint indicates a beneficial effect). The goal is to compute power as a function of the total sample size in the trial (total sample size ranges from 210 to 300).

Case study 7

This case study illustrates simulation-based power calculations in two-arm clinical trials with a superiority objective when the primary endpoint is a time-to-event endpoint. It is based on the same clinical trial for the treatment of metastatic colorectal cancer as Case study 8 for the TraditionalSampleSize module. An unbalanced design with 2:1 randomization (experimental treatment versus best supportive care) is utilized in the trial. The primary endpoint is the time to disease progression or death and is assumed to follow an exponential distribution. The primary analysis is performed using the log-rank test to show that the treatment is superior to the control. The goal is to compute power as a function of the target number of events (number of events ranges from 150 to 210) without accounting for loss to follow-up or administrative censoring, in which case the number of events is equal to the sample size (number of enrolled patients).

Case study 8

This case study illustrates simulation-based power calculations in two-arm clinical trials with a superiority objective when the primary endpoint is a time-to-event endpoint. It is based on the same clinical trial for the treatment of metastatic colorectal cancer as Case study 9 for the TraditionalSampleSize module. An unbalanced design with 2:1 randomization (experimental treatment versus best supportive care) is utilized in the trial. The primary endpoint is the time to disease progression or death and is assumed to follow an exponential distribution. The primary analysis is performed using the log-rank test to show that the treatment is superior to the control. The goal is to compute assurance as a function of the target number of events (number of events ranges from 150 to 210) without accounting for loss to follow-up or administrative censoring, in which case the number of events is equal to the sample size (number of enrolled patients). The assurance calculations are performed using the prior distributions for the hazard rates in the control and treatment arms that are derived from a historical trial with 50 events in the control arm and 40 events in the treatment arm.

Case study 9

This case study illustrates simulation-based power calculations in two-arm clinical trials with a non-inferiority objective when the primary endpoint is a time-to-event endpoint. It is based on the same clinical trial for the treatment of metastatic colorectal cancer as Case study 11 for the TraditionalSampleSize module. An unbalanced design with 2:1 randomization (experimental treatment versus best supportive care) is utilized in the trial. The primary endpoint is the time to disease progression or death and is assumed to follow an exponential distribution. The primary analysis is performed using the log-rank test to show that the treatment is non-inferior to the control. The non-inferiority margin on the hazard ratio scale is 1.3 (non-inferiority margin is greater than 1 since a larger value of the primary endpoint indicates a beneficial effect). The goal is to compute power as a function of the target number of events (number of events ranges from 600 to 690) without accounting for loss to follow-up or administrative censoring, in which case the number of events is equal to the sample size (number of enrolled patients).

Case study 10

This case study illustrates simulation-based power calculations in two-arm clinical trials with a superiority objective when the primary endpoint is a time-to-event endpoint. It is based on the same clinical trial for the treatment of metastatic colorectal cancer as Case study 12 for the TraditionalSampleSize module. An unbalanced design with 2:1 randomization (experimental treatment versus best supportive care) is utilized in the trial. The primary endpoint is the time to disease progression or death and is assumed to follow an exponential distribution. The primary analysis is performed using the log-rank test to show that the treatment is superior to the control. The trial includes a 9-month accrual period (patients are assumed to be recruited at a uniform rate) and the annual dropout rate is 5% (time to dropout is assumed to be exponentially distributed). The goal is to compute power as a function of the total number of enrolled patients (total number of patients ranges from 450 to 510) with the target number of events set to 380. The following important feature is illustrated in this case study: the user can specify a range of sample sizes across the samples in the data model and a single value for the target number of events. With this specification, multiple scenarios for the number of enrolled patients are considered but, across these scenarios, the patients will be followed until the fixed number of events is accrued. Also, since patients are followed up until the target number of events is met, the length of the follow-up period is not specified.

Case study 11

This case study illustrates simulation-based power calculations in two-arm clinical trials with a superiority objective when the primary endpoint is a time-to-event endpoint. It is based on the same clinical trial for the treatment of metastatic colorectal cancer as Case study 12 for the TraditionalSampleSize module. An unbalanced design with 2:1 randomization (experimental treatment versus best supportive care) is utilized in the trial. The primary endpoint is the time to disease progression or death and is assumed to follow an exponential distribution. The primary analysis is performed using the log-rank test to show that the treatment is superior to the control. The trial includes a 9-month accrual period (patients are assumed to be recruited at a uniform rate) and the annual dropout rate is 5% (time to dropout is assumed to be exponentially distributed). The goal is to compute power as a function of the total number of enrolled patients (total number of patients ranges from 450 to 510) with the target number of events ranging from 370 to 390. It is important to note that in this example, unlike Case study 10, multiple scenarios for the number of enrolled patients are considered and the target number of events is specified for each scenario. Since patients are followed up until the target number of events is met, the length of the follow-up period is not specified.

Case study 12

This case study illustrates simulation-based power calculations in two-arm clinical trials with a superiority objective when the primary endpoint is a time-to-event endpoint. It is based on the same clinical trial for the treatment of metastatic colorectal cancer as Case study 13 for the

TraditionalSampleSize module. An unbalanced design with 2:1 randomization (experimental treatment versus best supportive care) is utilized in the trial. The primary endpoint is the time to disease progression or death and is assumed to follow an exponential distribution. The primary analysis is performed using the log-rank test to show that the treatment is superior to the control. The trial includes a 9-month accrual period and the patient accrual is governed by a truncated exponential distribution with the median accrual time of 6 months. The annual dropout rate is 5% (time to dropout is assumed to be exponentially distributed). The goal is to compute power as a function of the total number of enrolled patients (total number of patients ranges from 450 to 510) with the target number of events ranging from 370 to 390. As in Case study 11, multiple scenarios for the number of enrolled patients are considered and the target number of events is specified for each scenario. Since patients are followed up until the target number of events is met, the length of the follow-up period is not specified.

Case study 13

This case study illustrates simulation-based power calculations in multi-arm clinical trials with a normally distributed primary endpoint. It is based on a Phase III clinical trial for the treatment of schizophrenia. Three doses of the experimental treatment (labeled Dose L, Dose M and Dose H) are tested versus placebo using a balanced design (1:1:1:1 randomization). The primary endpoint is the change in the Positive and Negative Syndrome Scale (PANSS) total score and a reduction in the total score is associated with improvement. This endpoint follows a normal distribution and each dose-placebo comparison is carried out using the two-sample Z-test to demonstrate that the dose is superior to placebo. The goal is to compute marginal power of each dose-placebo test as well as disjunctive power (probability of at least one significant dose-placebo test) as a function of the total sample size in the trial (total sample size ranges from 800 to 1040). To protect the overall Type I error rate in the trial, several candidate multiplicity adjustments are used (Bonferroni, Holm, Hochberg, Hommel and fixed-sequence procedures). The dose-placebo tests are assumed to be equally weighted when the Bonferroni, Holm, Hochberg and Hommel procedures are applied.

Case study 14

This case study illustrates simulation-based power calculations in multi-arm clinical trials with a normally distributed primary endpoint. It is based on a Phase III clinical trial for the treatment of schizophrenia. Three doses of the experimental treatment (labeled Dose L, Dose M and Dose H) are tested versus placebo using a balanced design (1:1:1:1 randomization). The primary endpoint is the change in the Positive and Negative Syndrome Scale (PANSS) total score and a reduction in the total score is associated with improvement. This endpoint follows a normal distribution and each dose-placebo comparison is carried out using the two-sample Z-test to demonstrate that the dose is superior to placebo. The goal is to compute marginal power of each dose-placebo test, disjunctive power (probability of at least one significant dose-placebo test) and weighted power with equal test weights as a function of the total sample size in the trial (total sample size ranges from 800 to 1040). The Hochberg procedure with several sets of test-specific weights is applied to protect the overall Type I error rate in the trial.

Case study 15

This case study illustrates simulation-based power calculations in multi-arm clinical trials with a normally distributed primary endpoint. It is based on a Phase III clinical trial for the treatment of schizophrenia. Three doses of the experimental treatment (labeled Dose L, Dose M and Dose H) are tested versus placebo using a balanced design (1:1:1:1 randomization). The primary endpoint is the change in the Positive and Negative Syndrome Scale (PANSS) total score and a reduction in the total score is associated with improvement. This endpoint follows a normal distribution and each dose-placebo comparison is carried out using the two-sample Z-test to demonstrate that the dose is superior to placebo. The goal is to compute marginal power of each dose-placebo test and disjunctive power (probability of at least one significant dose-placebo test) as a function of the total sample size in the trial (total sample size ranges from 800 to 1040). Several chain procedures are applied to protect the overall Type I error rate in the trial.

Case study 16

This case study illustrates simulation-based power calculations in multi-population clinical trials with a normally distributed primary endpoint. It is based on a Phase III clinical trial in patients with mild or moderate asthma. A two-arm balanced design with 1:1 randomization (experimental treatment versus placebo) is employed in the trial. The primary endpoint is defined as the change from baseline in the forced expiratory volume in one second (FEV1). This endpoint is normally distributed and a beneficial effect is associated with a larger increase in FEV1. The treatment effect is evaluated in the overall trial population (OP) as well as the pre-defined population of marker-positive patients (M+ population). Forty percent of the patients in the trial are expected to be marker-positive and the rest of the patients are marker-negative (M- population). The primary analysis in each patient population is carried out using the two-sample Z-test to demonstrate that the treatment is superior to placebo. The goal is to compute marginal power of the primary test within each patient population, disjunctive power (probability of a significant treatment effect within at least one population) and weighted power with unequal test weights as a function of the total sample size in the trial (total sample size ranges from 200 to 360) using several candidate multiplicity adjustments (Bonferroni, Hochberg and fixed-sequence procedures).

GroupSequential module

Case study 1

This case study illustrates analytical calculations in two-arm group-sequential trials when the primary endpoint is normally distributed. It is based on a Phase III clinical trial for the treatment of schizophrenia. A two-arm design (experimental treatment versus placebo) is utilized in the trial. The primary endpoint is the change in the Positive and Negative Syndrome Scale (PANSS) total score and a reduction in the total score indicates improvement. The primary endpoint follows a normal distribution and the primary analysis relies on the two-sample Z-test. The treatment effect is evaluated at a single interim analysis (at 50% of the total sample size) and at the final analysis. The efficacy boundary is defined using the O'Brien-Fleming spending function and the futility boundary is not specified. The goal is to perform analytical sample size calculations without modeling patient accrual or dropout.

Case study 2

This case study illustrates analytical and simulation-based calculations in two-arm group-sequential trials when the primary endpoint is normally distributed. It is based on a Phase III clinical trial for the treatment of schizophrenia. A two-arm design (experimental treatment versus placebo) is utilized in the trial. The primary endpoint is the change in the Positive and Negative Syndrome Scale (PANSS) total score and a reduction in the total score indicates improvement. The primary endpoint follows a normal distribution and the primary analysis relies on the two-sample Z-test. The treatment effect is evaluated at a single interim analysis (at 50% of the total sample size) and at the final analysis. The efficacy boundary is defined using the O'Brien-Fleming spending function and the futility boundary is not specified. The first goal is to perform analytical sample size calculations without modeling patient accrual or dropout. Secondly, a simulation-based approach is employed to compute key operating characteristics using the same treatment effect assumptions and assuming a 26-week accrual period with uniform patient recruitment, 6-week treatment period and a 25% dropout rate.

Case study 3

This case study illustrates analytical and simulation-based calculations in two-arm group-sequential trials when the primary endpoint is normally distributed. It is based on a Phase III clinical trial for the treatment of schizophrenia. A two-arm design (experimental treatment versus placebo) is utilized in the trial. The primary endpoint is the change in the Positive and Negative Syndrome Scale (PANSS) total score and a reduction in the total score indicates improvement. The primary endpoint follows a normal distribution and the primary analysis relies on the two-sample Z-test. The treatment effect is evaluated at two interim analyses (at 30% and 60% of the total sample size) and at the final analysis. The efficacy boundary is defined using the Pocock spending function and the futility boundary is not specified. The first goal is to perform analytical sample size calculations without modeling patient accrual or dropout. Secondly, a simulation-based approach is employed to compute key operating characteristics using the same treatment effect assumptions and assuming a 26-week accrual period with uniform patient recruitment, 6-week treatment period and a 25% dropout rate.

Case study 4

This case study illustrates analytical and simulation-based calculations in two-arm group-sequential trials when the primary endpoint is normally distributed. It is based on a Phase III clinical trial for the treatment of schizophrenia. A two-arm design (experimental treatment versus placebo) is utilized in the trial. The primary endpoint is the change in the Positive and Negative Syndrome Scale (PANSS) total score and a reduction in the total score indicates improvement. The primary endpoint follows a normal distribution and the primary analysis relies on the two-sample Z-test. The treatment effect is evaluated at two interim analyses (at 30% and 60% of the total sample size) and at the final analysis. The efficacy boundary is specified using the Hwang-Shih-DeCani spending function (gamma is set to -4 to define a boundary similar to the O'Brien-Fleming boundary) and the futility

boundary is not specified. The first goal is to perform analytical sample size calculations without modeling patient accrual or dropout. Secondly, a simulation-based approach is employed to compute key operating characteristics using the same treatment effect assumptions and assuming a 26-week accrual period with uniform patient recruitment, 6-week treatment period and a 25% dropout rate.

Case study 5

This case study illustrates analytical and simulation-based calculations in two-arm group-sequential trials when the primary endpoint is normally distributed. It is based on a Phase III clinical trial for the treatment of schizophrenia. A two-arm design (experimental treatment versus placebo) is utilized in the trial. The primary endpoint is the change in the Positive and Negative Syndrome Scale (PANSS) total score and a reduction in the total score indicates improvement. The primary endpoint follows a normal distribution and the primary analysis relies on the two-sample Z-test. The treatment effect is evaluated at a single interim analysis (at 50% of the total sample size) and at the final analysis. The efficacy boundary is defined using the O'Brien-Fleming spending function and the futility boundary is specified using the Hwang-Shih-DeCani spending function (gamma is set to 1 to define a boundary similar to the Pocock boundary). The first goal is to perform analytical sample size calculations without modeling patient accrual or dropout. Secondly, a simulation-based approach is employed to compute key operating characteristics using the same treatment effect assumptions and assuming a 26-week accrual period with uniform patient recruitment, 6-week treatment period and a 25% dropout rate.

Case study 6

This case study illustrates analytical and simulation-based calculations in two-arm group-sequential trials when the primary endpoint is normally distributed. It is based on a Phase III clinical trial for the treatment of schizophrenia. A two-arm design (experimental treatment versus placebo) is utilized in the trial. The primary endpoint is the change in the Positive and Negative Syndrome Scale (PANSS) total score and a reduction in the total score indicates improvement. The primary endpoint follows a normal distribution and the primary analysis relies on the two-sample Z-test. The treatment effect is evaluated at two interim analyses (at 30% and 60% of the total sample size) and at the final analysis. The efficacy boundary is specified using the Hwang-Shih-DeCani spending function (gamma is set to -4 to define a boundary similar to the O'Brien-Fleming boundary) and the futility boundary is also specified using the Hwang-Shih-DeCani spending function (gamma is set to 1 to define a boundary similar to the Pocock boundary). The first goal is to perform analytical sample size calculations without modeling patient accrual or dropout. Secondly, a simulation-based approach is employed to compute key operating characteristics using the same treatment effect assumptions and assuming a 26-week accrual period with uniform patient recruitment, 6-week treatment period and a 25% dropout rate.

Case study 7

This case study illustrates analytical and simulation-based calculations in two-arm group-sequential trials when the primary endpoint is normally distributed. It is based on a Phase III clinical trial for the treatment of schizophrenia. A two-arm design (experimental treatment versus placebo) is utilized in the trial. The primary endpoint is the change in the Positive and Negative Syndrome Scale (PANSS) total score and a reduction in the total score indicates improvement. The primary endpoint follows a normal distribution and the primary analysis relies on the two-sample Z-test. The treatment effect is evaluated at two interim analyses (at 30% and 60% of the total sample size) and at the final analysis. The efficacy boundary is specified using the Hwang-Shih-DeCani spending function (gamma is set to -4 to define a boundary similar to the O'Brien-Fleming boundary) and the futility boundary is also specified using the Hwang-Shih-DeCani spending function (gamma is set to 1 to define a boundary similar to the Pocock boundary). The first goal is to perform analytical sample size calculations without modeling patient accrual or dropout. Secondly, a simulation-based approach is employed to compute key operating characteristics using the same treatment effect assumptions and assuming a 26-week accrual period (patient accrual is governed by a truncated exponential

distribution with the median accrual time of 18 weeks), 6-week treatment period and a 25% dropout rate.

Case study 8

This case study illustrates analytical and simulation-based calculations in two-arm group-sequential trials when the primary endpoint is normally distributed. It is based on a Phase III clinical trial for the treatment of schizophrenia. A two-arm design (experimental treatment versus placebo) is utilized in the trial. The primary endpoint is the change in the Positive and Negative Syndrome Scale (PANSS) total score and a reduction in the total score indicates improvement. The primary endpoint follows a normal distribution and the primary analysis relies on the two-sample Z-test. The treatment effect is evaluated at a single interim analysis (at 50% of the total sample size) and at the final analysis. The efficacy boundary is defined using the O'Brien-Fleming spending function and the futility boundary is not specified. The first goal is to perform analytical sample size calculations without modeling patient accrual or dropout. Secondly, a simulation-based approach is employed to compute key operating characteristics using an alternative set of treatment effect assumptions (a smaller treatment difference) and assuming a 26-week accrual period with uniform patient recruitment, 6-week treatment period and a 25% dropout rate.

Case study 9

This case study illustrates analytical and simulation-based calculations in two-arm group-sequential trials when the primary endpoint is normally distributed. It is based on a Phase III clinical trial for the treatment of schizophrenia. A two-arm design (experimental treatment versus placebo) is utilized in the trial. The primary endpoint is the change in the Positive and Negative Syndrome Scale (PANSS) total score and a reduction in the total score indicates improvement. The primary endpoint follows a normal distribution and the primary analysis relies on the two-sample Z-test. The treatment effect is evaluated at a single interim analysis (at 50% of the total sample size) and at the final analysis. The efficacy boundary is defined using the O'Brien-Fleming spending function and the futility boundary is not specified. The first goal is to perform analytical sample size calculations without modeling patient accrual or dropout. Secondly, a simulation-based approach is employed to compute key operating characteristics assuming a 26-week accrual period with uniform patient recruitment, 6-week treatment period and a 25% dropout rate. The simulation-based approach focuses on assurance calculations that are performed using fairly vague prior distributions for the means and standard deviations in the placebo and treatment arms. These priors are derived from a small historical trial with 50 patients in each arm.

Case study 10

This case study illustrates analytical calculations in two-arm group-sequential trials when the primary endpoint is binary. It is based on a Phase III clinical trial in patients with rheumatoid arthritis. A two-arm design (experimental treatment versus placebo) is utilized in the trial. The primary endpoint is based on the American College of Rheumatology (ACR50) definition of improvement (a higher response rate indicates a beneficial treatment effect) and the primary analysis relies on the two-sample Z-test for proportions. The treatment effect is evaluated at a single interim analysis (at 50% of the total sample size) and at the final analysis. The efficacy boundary is defined using the O'Brien-Fleming spending function and the futility boundary is not specified. The first goal is to perform analytical sample size calculations without modeling patient accrual or dropout.

Case study 11

This case study illustrates analytical and simulation-based calculations in two-arm group-sequential trials when the primary endpoint is binary. It is based on a Phase III clinical trial in patients with rheumatoid arthritis. A two-arm design (experimental treatment versus placebo) is utilized in the trial. The primary endpoint is based on the American College of Rheumatology (ACR50) definition of improvement (a higher response rate indicates a beneficial treatment effect) and the primary analysis relies on the two-sample Z-test for proportions. The treatment effect is evaluated at two

interim analyses (at 30% and 60% of the total sample size) and at the final analysis. The efficacy boundary is defined using the Pocock spending function and the futility boundary is not specified. The first goal is to perform analytical sample size calculations without modeling patient accrual or dropout. Secondly, a simulation-based approach is employed to compute key operating characteristics using the same treatment effect assumptions and assuming a 12-month accrual period with uniform patient recruitment, 6-month treatment period and a 10% dropout rate.

Case study 12

This case study illustrates analytical and simulation-based calculations in two-arm group-sequential trials when the primary endpoint is binary. It is based on a Phase III clinical trial in patients with rheumatoid arthritis. A two-arm design (experimental treatment versus placebo) is utilized in the trial. The primary endpoint is based on the American College of Rheumatology (ACR50) definition of improvement (a higher response rate indicates a beneficial treatment effect) and the primary analysis relies on the two-sample Z-test for proportions. The treatment effect is evaluated at two interim analyses (at 30% and 60% of the total sample size) and at the final analysis. The efficacy boundary is specified using the Hwang-Shih-DeCani spending function (gamma is set to -4 to define a boundary similar to the O'Brien-Fleming boundary) and the futility boundary is also specified using the Hwang-Shih-DeCani spending function (gamma is set to 1 to define a boundary similar to the Pocock boundary). The first goal is to perform analytical sample size calculations without modeling patient accrual or dropout. Secondly, a simulation-based approach is employed to compute key operating characteristics using the same treatment effect assumptions and assuming a 12-month accrual period (patient accrual is governed by a truncated exponential distribution with the median accrual time of 9 months), 6-month treatment period and a 10% dropout rate.

Case study 13

This case study illustrates analytical and simulation-based calculations in two-arm group-sequential trials when the primary endpoint is binary. It is based on a Phase III clinical trial in patients with rheumatoid arthritis. A two-arm design (experimental treatment versus placebo) is utilized in the trial. The primary endpoint is based on the American College of Rheumatology (ACR50) definition of improvement (a higher response rate indicates a beneficial treatment effect) and the primary analysis relies on the two-sample Z-test for proportions. The treatment effect is evaluated at a single interim analysis (at 50% of the total sample size) and at the final analysis. The efficacy boundary is defined using the O'Brien-Fleming spending function and the futility boundary is not specified. The first goal is to perform analytical sample size calculations without modeling patient accrual or dropout. Secondly, a simulation-based approach is employed to compute key operating characteristics using an alternative set of treatment effect assumptions (a smaller treatment difference) and assuming a 12-month accrual period with uniform patient recruitment, 6-month treatment period and a 10% dropout rate.

Case study 14

This case study illustrates analytical and simulation-based calculations in two-arm group-sequential trials when the primary endpoint is binary. It is based on a Phase III clinical trial in patients with rheumatoid arthritis. A two-arm design (experimental treatment versus placebo) is utilized in the trial. The primary endpoint is based on the American College of Rheumatology (ACR50) definition of improvement (a higher response rate indicates a beneficial treatment effect) and the primary analysis relies on the two-sample Z-test for proportions. The treatment effect is evaluated at a single interim analysis (at 50% of the total sample size) and at the final analysis. The efficacy boundary is defined using the O'Brien-Fleming spending function and the futility boundary is not specified. The first goal is to perform analytical sample size calculations without modeling patient accrual or dropout. Secondly, a simulation-based approach is employed to compute key operating characteristics assuming a 12-month accrual period with uniform patient recruitment, 6-month treatment period and a 10% dropout rate. The simulation-based approach focuses on assurance calculations that are performed using fairly vague prior distributions for the response rates in the placebo and treatment arms. These priors are derived from a historical trial with 80 patients in each arm.

Case study 15

This case study illustrates analytical calculations in two-arm group-sequential trials with a time-to-event primary endpoint. It is based on a Phase III clinical trial in patients with metastatic colorectal cancer. A two-arm design (experimental treatment versus best supportive care) is utilized in the trial. The primary endpoint is the time to disease progression or death and is assumed to follow an exponential distribution. The treatment effect is evaluated using the log-rank test. The treatment effect is evaluated at a single interim analysis (at 50% of the total sample size) and at the final analysis. The efficacy boundary is defined using the O'Brien-Fleming spending function and the futility boundary is not specified. The goal is to perform analytical sample size calculations without modeling patient accrual or dropout.

Case study 16

This case study illustrates analytical and simulation-based calculations in two-arm group-sequential trials with a time-to-event primary endpoint. It is based on a Phase III clinical trial in patients with metastatic colorectal cancer. A two-arm design (experimental treatment versus best supportive care) is utilized in the trial. The primary endpoint is the time to disease progression or death and is assumed to follow an exponential distribution. The treatment effect is evaluated using the log-rank test. The treatment effect is evaluated at a single interim analysis (at 50% of the total sample size) and at the final analysis. The efficacy boundary is defined using the O'Brien-Fleming spending function and the futility boundary is not specified. The first goal is to perform analytical sample size calculations without modeling patient accrual or dropout. Secondly, a simulation-based approach is employed to compute key operating characteristics using the same treatment effect assumptions and assuming a 9-month accrual period with uniform patient recruitment and an annual dropout rate of 5%. The simulations are performed under the assumptions that 250 patients per arm are enrolled into the trial (these patients are followed up until the target number of events is met and the length of the follow-up period is not specified).

Case study 17

This case study illustrates analytical and simulation-based calculations in two-arm group-sequential trials with a time-to-event primary endpoint. It is based on a Phase III clinical trial in patients with metastatic colorectal cancer. A two-arm design (experimental treatment versus best supportive care) is utilized in the trial. The primary endpoint is the time to disease progression or death and is assumed to follow an exponential distribution. The treatment effect is evaluated using the log-rank test. The treatment effect is evaluated at a single interim analysis (at 50% of the total sample size) and at the final analysis. The efficacy boundary is specified using the Hwang-Shih-DeCani spending function (gamma is set to -4 to define a boundary similar to the O'Brien-Fleming boundary) and the futility boundary is also specified using the Hwang-Shih-DeCani spending function (gamma is set to 1 to define a boundary similar to the Pocock boundary). The first goal is to perform analytical sample size calculations without modeling patient accrual or dropout. Secondly, a simulation-based approach is employed to compute key operating characteristics using the same treatment effect assumptions and assuming a 9-month accrual period with uniform patient recruitment and an annual dropout rate of 5%. The simulations are performed under the assumptions that 250 patients per arm are enrolled into the trial (these patients are followed up until the target number of events is met and the length of the follow-up period is not specified).

Case study 18

This case study illustrates analytical and simulation-based calculations in two-arm group-sequential trials with a time-to-event primary endpoint. It is based on a Phase III clinical trial in patients with metastatic colorectal cancer. A two-arm design (experimental treatment versus best supportive care) is utilized in the trial. The primary endpoint is the time to disease progression or death and is assumed to follow an exponential distribution. The treatment effect is evaluated using the log-rank test. The treatment effect is evaluated at a single interim analysis (at 50% of the total sample size) and at the final analysis. The efficacy boundary is specified using the Hwang-Shih-DeCani spending function (gamma is set to -4 to define a boundary similar to the O'Brien-Fleming boundary) and the

futility boundary is also specified using the Hwang-Shih-DeCani spending function (gamma is set to 1 to define a boundary similar to the Pocock boundary). The first goal is to perform analytical sample size calculations without modeling patient accrual or dropout. Secondly, a simulation-based approach is employed to compute key operating characteristics using the same treatment effect assumptions and assuming a 9-month accrual period (patient accrual is governed by a truncated exponential distribution with the median accrual time of 6 months) and an annual dropout rate of 5%. The simulations are performed under the assumptions that 250 patients per arm are enrolled into the trial (these patients are followed up until the target number of events is met and the length of the follow-up period is not specified).

Case study 19

This case study illustrates analytical and simulation-based calculations in two-arm group-sequential trials with a time-to-event primary endpoint. It is based on a Phase III clinical trial in patients with metastatic colorectal cancer. A two-arm design (experimental treatment versus best supportive care) is utilized in the trial. The primary endpoint is the time to disease progression or death and is assumed to follow an exponential distribution. The treatment effect is evaluated using the log-rank test. The treatment effect is evaluated at a single interim analysis (at 50% of the total sample size) and at the final analysis. The efficacy boundary is specified using the Hwang-Shih-DeCani spending function (gamma is set to -4 to define a boundary similar to the O'Brien-Fleming boundary) and the futility boundary is also specified using the Hwang-Shih-DeCani spending function (gamma is set to 1 to define a boundary similar to the Pocock boundary). The first goal is to perform analytical sample size calculations without modeling patient accrual or dropout. Secondly, a simulation-based approach is employed to compute key operating characteristics using an alternative set of treatment effect assumptions (a weaker treatment effect) and assuming a 9-month accrual period (patient accrual is governed by a truncated exponential distribution with the median accrual time of 6 months) and an annual dropout rate of 5%. The simulations are performed under the assumptions that 250 patients per arm are enrolled into the trial (these patients are followed up until the target number of events is met and the length of the follow-up period is not specified).

Case study 20

This case study illustrates analytical calculations in two-arm group-sequential trials with a time-to-event primary endpoint. It is based on a Phase III clinical trial in patients with metastatic colorectal cancer. A two-arm design (experimental treatment versus best supportive care) is utilized in the trial. The primary endpoint is the time to disease progression or death and is assumed to follow an exponential distribution. The treatment effect is evaluated using the log-rank test. The treatment effect is evaluated at a single interim analysis (at 50% of the total sample size) and at the final analysis. The efficacy boundary is specified using the Hwang-Shih-DeCani spending function (gamma is set to -4 to define a boundary similar to the O'Brien-Fleming boundary) and the futility boundary is also specified using the Hwang-Shih-DeCani spending function (gamma is set to 1 to define a boundary similar to the Pocock boundary). The first goal is to perform analytical sample size calculations without modeling patient accrual or dropout. Secondly, a simulation-based approach is employed to compute key operating characteristics using the same treatment effect assumptions and assuming a 9-month accrual period (patient accrual is governed by a truncated exponential distribution with the median accrual time of 6 months) and an annual dropout rate of 5%. The simulations are performed under the assumptions that 250 patients per arm are enrolled into the trial (these patients are followed up until the target number of events is met and the length of the follow-up period is not specified). Also, the simulation-based approach focuses on assurance calculations that are performed using prior distributions for the hazard rates in the placebo and treatment arms that are derived from a smaller historical trial (90 events in the control arm and 70 events in the treatment arm).