Sample sizes for non-inferiority studies based on the difference between two proportions: a unified approach for difference, ratio and odds ratio models.

Bruno Mario Cesana [1,2], Paolo Antonelli [1]

(1) Unit of Biostatistics and Biomathematics, Department of Molecular and Transactional Medicine, University of Brescia, Brescia, Italy
(2) Department of Clinical Sciences and Community Health, Unit of Medical Statistics, Biometry and Bioinformatics "Giulio A. Maccacaro", Faculty of Medicine and Surgery, University of Milan, Milan, Italy

CORRESPONDING AUTHOR: *Bruno Mario Cesana,: Department of Clinical Sciences and Community Health, Unit of Medical Statistics, Biometry and Bioinformatics "Giulio A. Maccacaro", Faculty of Medicine and Surgery, University of Milan, Milan, Italy. Campus Cascina Rosa, Fondazione IRCCS Istituto Nazionale Tumori, Via A. Vanzetti 5, 20133 Milano, Italy. E-mail: brnmrcesana@gmail.com*

DOI: ***10-2427/13265***
Accepted on February 10, 2020

ABSTRACT

Background: It is never easy to make sample size calculation for two-arm, non-inferiority studies with a primary proportion outcome not only because the different parameterisations involved in the difference, ratio, and odds ratio models lead to different non-inferiority margins and different sample size results, but also because of the different efficiency of the respective sample size formulae.
Methods: According to a formal statistical approach, we showed how to express the non-inferiority margins of the three models by keeping the probability (success or failure) of the standard treatment fixed (considered as "known" in the planning phase of a trail), and equal under the null and alternative hypotheses as the statistical basis for sample size calculation.
Results: We have obtained the sample size formulae and their respective power formulae for the three considered models both for success and failure probabilities. A sample size table for non-inferiority success studies is reported for illustrative purposes. In addition, we have compared the sample sizes from the three models by means of graphic and theoretical approaches and we have shown their asymptotic relationships.
Furthermore, we have obtained the formulae for switching among the three considered models. Finally, we have correct some previously published formulae for sample size calculations.
Conclusion: The clearly separate approach to the probabilities of success and failure of the three considered models shown in this paper makes it possible to switch among them consistently and equivalently and to choose the probability formulation for the most parsimonious model.

Key words: Clinical trials, Sample size calculation, Non-inferiority studies, Success and failure probabilities, Difference, ratio, odds ratio models

1. INTRODUCTION

Non-inferiority controlled clinical trials having the proportion of success or failure as the primary outcome are increasingly being carried out particularly in the fields of cardiology, oncology and antibiotics and their ethical nature [1,2] nowadays has to be taken for granted.

The opposite roles of the null (H_0) and alternative hypotheses (H_A) makes their rationale a little counter-intuitive and their statistical testing not easy as it has been showed firstly by Dunnett and Gent [3], as a formal statistical significance test or from the exact confidence intervals of the odds ratio according to Gart [4].

Furthermore, the choice of the maximum difference that is not clinically/biologically relevant leading to an Experimental drug being considered "non-inferior" (i.e. the Non-Inferiority Margin: NIM) is central to the scientific and ethical plausibility of a non-inferiority trial and the validity of its conclusions.

We cannot consider in detail the suggestions about the choice of the NIM given by regulatory guidelines [5,6,7,8,9], the discrepancies between the FDA and the European Medicines Agency (EMA) guidelines for trials in diabetes mellitus [10,11] and in infectious diseases [12,13,14] together with the methodological attempt to reconcile them by Röhmel [15], and the several proposals, based on a percentage of the expected difference between the standard and the placebo, suggested, among others, by Holmgren [16], by D'Agostino et al. [17] by Snapinn [18] or by Pigeot et al. [19]. Finally, further recent insights are in Hung et al. [20,21], Wiens [22], and Tsong et al. [23].

At least, it has to be stated that the choice of the NIM has to be based on clinical and statistical criteria, that the NIM has to be lower than the smallest difference between the Standard and placebo, and the NIM has to be justified, on the fact that the Experimental (which is expected to be non-inferior) clearly has a real advantage over the Standard (easier administration, fewer adverse events because of its biologically well-documented mechanism, lower cost, etc.) [24].

In addition, the choice of the NIM in two arms trials (without a placebo for ethical reasons) has to fulfil "assay sensitivity" (the Experimental is efficacious in the sense that it would be superior to placebo or the previous standard) and the "constancy assumption" (the Standard effect remains the same) [17].

A second aspect is the parameterisation of the NIM. The odds ratio scale has been supported by Julious [25], Garrett [26], Senn [27], Tu [28], Siegel [29], Kaul and Diamond [30] by Wang et al.[31], by Chow et al.[32] who showed some sample size calculations for equality, non-inferiority/superiority and equivalence trials in the context of parallel and cross-over designs, and, finally, by the FDA guideline [8], but only in the case of a lower event rate, or when the reference treatment is expected to have response rates near 0% or 100% according to the CPMP guideline [6].

However, it is necessary to consider the impact of odds ratio parameterisation on sample size calculation and, in the words of PASS®[33] "As a rule of thumb, the difference is best suited for those cases in which $0.20<P<0.80$", taking also into account that this scale is more familiar to clinicians.

In this paper, we consider the parameterisation of the difference (D, the most familiar to clinicians), the ratio (R), the natural logarithm of the ratio [LR], and the odds ratio (OR) or, better, the natural logarithm of the odds ratio [LOR].

Furthermore, as a third relevant point there is still no agreed approach to sample size calculation for the case of the difference between two proportions.

Apart from the papers of Makuch and Simon [34], Blackwelder [35] dealing with the sample size calculation for the comparison of two proportions of success in non-inferiority studies and the papers of Blackwelder and Chang [36], Heiselbetz and Edler [37] providing graphs of the sample size and a computer program, respectively, we want to draw the attention on the Farrington and Manning's paper [38] owing to its relevance.

Indeed, Farrington and Manning [38] considered three methods of obtaining the approximate variance of the difference between two proportions under the null hypothesis of a non-zero difference; i) the "observed values" of Dunnett and Gent [3], Makuch and Simon [34], and Blackwelder [35] (method 1); ii) the values obtained from the "fixed marginal totals" of Dunnett and Gent [3] and Rodary et al.[39] (method 2); and iii) the values obtained using the "maximum likelihood estimation" such as the solutions of a cubic equation according to Miettinen and Nurminen [40], but disregarding the term (N1-1)/N1, which is negligible in large samples (method 3). It has to be pointed out that method 3 has also been proposed as a means of overcoming some of the serious drawbacks of the first two methods: the poor coverage of method 1, and the fact that the values obtained from the constrained estimation under H_0 of method 2 have to satisfy some easily violated inequalities [38]. Farrington and Manning [38] also showed the sample size calculations for a non-unity relative risk using the three methods, and included a sample size imbalance that is equal to the ratio between the sample sizes of the Experimental and the Standard.

However, in Farrington and Manning's paper [38] there is not clear distinction between the probabilities of success or failure (being positive or negative the non-inferiority margin of the difference and more or less than 1 the non-inferiority margin of the ratio), some of the considered scenarios are better suited to the H_0 formulation of a "superiority test" according to Chow et al.[32] and the sample sizes shown in Tables I and II [38] are not reproduced by using the PASS® software [41,42,43] since it is not pertinent to calculate the sample size in the case of the probability of the Experimental greater

than that of the Standard (compatible situation with a failure probability) with a non-inferiority margin negative instead of positive and the same applies to some cases with a ratio<1 instead of >1. Finally, the sample size allocation ratio of 3/2, even if it was chosen for illustrative purposes, it is misleading leading to an unusual situation of a larger number of patients in the standard group.

Farrington and Manning [38] did not compare the sample size results of the difference and ratio parameterisations of H_0. However, as can be seen from their Table II, method 1 almost always requires a larger sample size for the ratio than methods 2 and 3, which give similar sample sizes; in the case of the difference, method 3 is generally the most demanding in terms of sample size, with methods 1 and 2 (when feasible) being similar to each other.

Farrington & Manning's approach [38] has been implemented by Machin et al.[44,45,46] instead of the Makuch and Simon's approach [34] previously adopted [47]; however, the sample size calculation Formula 5.4 [44] and Formula 9.10 [45] and the sample size values reported in Tables 5.1[44] and 9.3 [45] are correct only in the trivial but most frequent case of an equal allocation ratio (see Appendix 1, paragraph 1.1). However, the above formulae have been corrected in the last (4th) edition of Machin et al.'s book [46].

Furthermore, Laster and Johnson first considered sample size calculations for non-inferiority trials with quantitative outcomes [48], and then, together with Kotler, for the ratio between Experimental and Standard probabilities [49]. They compared the results obtained from the parameterisation of the ratio with those obtained from the parameterisation of the difference as proposed by Blackwelder [35] and showed the conversion formula from the H_0 of a ratio into that of a difference for both success and failure probabilities. However, in the case of success probability, the conversion formula is sensibly based on the probability of the Standard since it has to be considered as absolutely known at the sample size calculation step: whereas, in the case of failure probability, the conversion formula is based on the probability of the Experimental according to a not shareable approach, in our opinion, since the NIM depends on the probability of the Standard. Then, by inverting the ratio (the Standard divided by the Experimental probability), they obtain a ratio of <1, thus leading to sample sizes that are smaller than those obtained for the difference parameterisation (the problems related to this questionable approach is discussed in Appendix 1, 1.2 paragraph). Furthermore, their sample size tables (Table I for success probability and Table II for failure probability) show ratio values of >1 between the true probabilities, thus leading to an expected superiority of the Experimental in the case of success probabilities, which is, obviously, not suitable for a non-inferiority trial. The same is true for ratio values of <1 in the case of failure probabilities.

Other papers on particular aspects of non-inferiority sample size calculations that are worth mentioning include those of Julious and Owen [50] even if the formulae attributed to Farrington and Manning's [38] maximum likelihood estimates and the confidence interval of Wilson's score [51] are not correct, leading to much larger sample sizes, Julious' book [25] in which it is used the Blackwelder's approach [35], Chan [52], Röhmel and Mansmann [53], Chan and Zhan [54], Kang and Chen [55], Chan [56,57] with exact methods for calculating exact significance levels, de Boo and Zielhuis [58] with their approach for obtaining the smallest total sample sizes when using the exact method for unequal sample sizes in both risk difference and relative risk situations, and, finally in the context of failure probability, Hilton [59,60] provided an algorithm for identifying the optimal imbalance between experimental and standard groups in order to obtain a minimum total sample size, although it has to be pointed out that a standard ratio of, for example, 2:1 or 3:2 is preferable.

However, there are no published comparisons of switching among the three models of difference (D), ratio (R) with its logarithm (LR), AND odds ratio (OR) with its logarithm (LOR), taking into account the different parameterisations of their statistical hypotheses and non-inferiority margins.

The aim of this paper is to show how to switch among the models of the difference (D), ratio of two probabilities (R), logarithm of the ratio [LR], and the log transformation of the odds ratio [LOR] for the probabilities of success and failure, and to switch from the probability of success to the probability of failure for the sample size calculation and the power of statistical significance tests.

In accordance with general methodology, we have consistently and equivalently formulated the different NIMs pertinent to the parameterisation of the models by starting with the probability of the success or failure of the standard treatment, which is assumed to be known during the planning phase of a study and, unlike the probability of the experimental treatment, is independent of the non-inferiority margin fixed by the researcher. We have used the same approach to switch between the probabilities of success and failure. Then, we have compared the different sample sizes calculated for the considered models in order to choose the most parsimonious.

2. METHODOLOGICAL BACKGROUND

2.A. Success probability

Let us consider two independently binomially distributed probabilities, with π_{S_St} and π_{S_Ex} as the true success/

favourable outcome parameters: the subscripts "S", "St" and "Ex" respectively stand for Success, Standard treatment and the Experimental new treatment, which is expected to be "not inferior" to the Standard.

$$p_{S\_St} \sim Bin\left(\pi_{S\_St}, n_{S\_St}\right) \text{ with } \sigma^2_{S\_St} = \frac{\pi_{S\_St}\left(1-\pi_{S\_St}\right)}{n_{S\_St}}$$

$$p_{S\_Ex} \sim Bin\left(\pi_{S\_Ex}, n_{S\_Ex}\right) \text{ with } \sigma^2_{S\_Ex} = \frac{\pi_{S\_Ex}\left(1-\pi_{S\_Ex}\right)}{n_{S\_Ex}}$$

In order to adopt a unified approach to the statistical significance test and sample size calculation, let us define T_S the statistic of interest: i.e. the difference between two success proportions (D), their ratio (R) and its natural logarithm (LR), their odds ratio (OR) and its natural logarithm (LOR), and assume that the distribution of T_S can, under suitable conditions, be approximated to a Gaussian distribution with mean value (μ_{T_S}) and variance ($\sigma^2_{T_S}$). Finally, as the expected value and the variance of T_S are different under H_0 and H_A, the subscript "H_0" and "H_A" will be used: i.e. $\mu_{T_S\_H_0}$ and $\sigma^2_{T_S\_H_0}$ under H_0, and $\mu_{T_S\_H_A}$ and $\sigma^2_{T_S\_H_A}$ under H_A.

2.A.1 Formulation of the H_0 and H_A hypotheses

With θ_S the parameter of interest and θ_{0_S} the maximal clinically/biologically irrelevant threshold (the non-inferiority margin), the non-inferiority hypotheses are:

$$H_0 : \theta_S \leq \theta_{0\_s} \text{ for inferiority}$$

$$H_A : \theta_S > \theta_{0\_s} \text{ for non-inferiority}$$

2.A.2 Statistical significance test

Given the above formulation of H_0, the non-inferiority statistical test will always be one-sided (on the right) with an approximate test function given by:

$$\frac{T_s - \mu_{T_S\_H_0}}{\sigma_{T_S\_H_0}} \approx Z(0;1) \qquad (2.A.2)$$

Defining t_S as the sampling value of T_S, and with a significance level of $\alpha = 0.05$ two-sided (or, equivalently, 0.025 one-sided), H_0 will be rejected if $z > z_{1-\alpha/2}$ or $t_S > t_c$, where t_c is the quantile delimiting the critical region: $t_c = \mu_{T_S\_H_0} + z_{1-\alpha/2}\sigma_{T_S\_H_0}$. However, usually, the non-inferiority H_0 hypothesis is rejected if the lower limit of the 95% confidence interval of the difference standard minus experimental is greater than the positive non-inferiority margin. The rejection of the null hypothesis would make the non-inferiority of the Experimental the most plausible conclusion.

It should be mentioned, in accordance with the definition that H_0 is greater than or equal to the non-inferiority margin, that statistical significance occurs when the test statistic is greater than the corresponding quantile of the Z distribution; however, in the case of continuous variables, this clarification is pragmatically irrelevant.

2.A.3 Sample size calculation

The rationale underlying the sample size calculation requires the simultaneous occurrence of two events: a statistically significant result (under H_0) and the rejection of H_0 under H_A (defined as the power of the test):

$$P\left\{T_S > t_c | H_0\right\} < \alpha/2 \text{ and } P\left\{T_S > t_c | H_A\right\} \geq 1-\beta$$

Solving the basic inequalities for t_c gives:

$$\alpha/2 > P(T_S > t_c \mid H_0) = P\left(Z > \frac{t_c - \mu_{T_S\_H_0}}{\sigma_{T_S\_H_0}}\right) = P(Z > z_{1-\alpha/2})$$

$$then: \quad \frac{t_c - \mu_{T_S\_H_0}}{\sigma_{T_S\_H_0}} \geq z_{1-\alpha/2} \rightarrow t_c \geq \mu_{T_S\_H_0} + z_{1-\alpha/2}\sigma_{T_S\_H_0}$$

and:

$$1-\beta \leq P(T_S > t_c \mid H_A) = P\left(Z > \frac{t_c - \mu_{T_S\_H_A}}{\sigma_{T_S\_H_A}}\right) = P(Z > z_\beta) = P(Z > -z_{1-\beta})$$

$$then: \quad \frac{t_c - \mu_{T_S\_H_A}}{\sigma_{T_S\_H_A}} \leq -z_{1-\beta} \rightarrow t_c \geq \mu_{T_S\_H_A} - z_{1-\beta}\sigma_{T_S\_H_A}$$

Note that z_β has been replaced by $-z_{1-\beta}$ because of the symmetry of the Z distribution.

Finally, by equating the above expressions to t_c, the sample size can be calculated using the following general pivotal formula:

$$z_{1-\alpha/2}\sigma_{T_S\_H_0} + z_{1-\beta}\sigma_{T_S\_H_A} = \mu_{T_S\_H_A} - \mu_{T_S\_H_0} \qquad (2.A.3)$$

which has to be explicitly solved for the sample sizes of the Experimental (n_{S_Ex}) and the Standard (n_{S_St}).

After having put $n_{S\_St} = k \cdot n_{S\_Ex}$, the following quantities can be defined in a conveniently simplified form (n_{S_Ex} is explicitly included because it is the solution of the sample size formula):

$$V^2_{T_S\_H_0} = n_{S\_Ex}\sigma^2_{T_S\_H_0}; \; V^2_{T_S\_H_A} = n_{S\_Ex}\sigma^2_{T_S\_H_A}$$

In this way, the general equation becomes:

$$z_{1-\alpha/2}\sqrt{\sigma^2_{T_S\_H_0}} + z_{1-\beta}\sqrt{\sigma^2_{T_S\_H_A}} = \mu_{T_S\_H_A} - \mu_{T_S\_H_0} \rightarrow z_{1-\alpha/2}\sqrt{\frac{1}{n_{S\_Ex}}V^2_{T_S\_H_0}} + z_{1-\beta}\sqrt{\frac{1}{n_{S\_Ex}}V^2_{T_S\_H_A}} = \mu_{T_S\_H_A} - \mu_{T_S\_H_0}$$

$$\rightarrow \frac{1}{\sqrt{n_{S\_Ex}}}\left(z_{1-\alpha/2}\sqrt{V^2_{T_S\_H_0}} + z_{1-\beta}\sqrt{V^2_{T_S\_H_A}}\right) = \mu_{T_S\_H_A} - \mu_{T_S\_H_0}$$

which, resolved by n_{S_Ex}, gives the general formula for the sample size calculation of the Experimental:

$$n_{S\_Ex} = \frac{\left(z_{1-\alpha/2}\sqrt{V^2_{T_S\_H_0}} + z_{1-\beta}\sqrt{V^2_{T_S\_H_A}}\right)^2}{\left(\mu_{T_S\_H_A} - \mu_{T_S\_H_0}\right)^2} \qquad (2.A.3.1)$$

This formula has to be appropriately adapted to the parameters of the considered models and, in order to allow for an unequal allocation, the ratio of the two sample sizes ($k = n_{S\_St} / n_{S\_Ex}$) has to be calculated, with n_{S_St} being calculated as $k \cdot n_{S\_Ex}$ rather than by using an ad hoc equation similar to 2.A.3.1.

2.A.4 Power of the statistical significance test

Given that $t_c = \mu_{T_S\_H_0} + z_{1-\alpha/2}\sigma_{T_S\_H_0}$, the power ($1-\beta$) of the statistical test under H_A is given by:

$$1-\beta = P(T_S \geq t_c \mid H_A) = P\left(Z \geq \frac{t_c - \mu_{T_S\_H_A}}{\sigma_{T_S\_H_A}}\right) = P\left(Z \geq \frac{\mu_{T_S\_H_0} + z_{1-\alpha}\sigma_{T_S\_H_0} - \mu_{T_S\_H_A}}{\sigma_{T_S\_H_A}}\right) =$$
$$= P\left(Z \geq \frac{z_{1-\alpha/2}\sigma_{T_S\_H_0} - \left(\mu_{T_S\_H_0} - \mu_{T_S\_H_0}\right)}{\sigma_{T_S\_H_A}}\right) = 1 - \Phi\left(\frac{z_{1-\alpha/2}\sigma_{T_S\_H_0} - \left(\mu_{T_S\_H_A} - \mu_{T_S\_H_0}\right)}{\sigma_{T_S\_H_A}}\right) = \quad (2.A.4.1)$$
$$= \Phi\left(\frac{\left(\mu_{T_S\_H_A} - \mu_{T_S\_H_0}\right) - z_{1-\alpha/2}\sigma_{T_S\_H_0}}{\sigma_{T_S\_H_A}}\right)$$

In addition, the power of the test can be more intuitively obtained by means of the sample size calculation formula resolved by $z_{1-\beta}$, and then by calculating its corresponding probability value ($1-\beta$):

$$z_{1-\beta}\sqrt{V^2_{T_S\_H_A}} = \left(\mu_{T_S\_H_A} - \mu_{T_S\_H_0}\right)\sqrt{n_{S\_E0}} - z_{1-\alpha/2}\sqrt{V^2_{T_S\_H_0}} \rightarrow$$
$$z_{1-\beta} = \frac{\left(\mu_{T_S\_H_A} - \mu_{T_S\_H_0}\right)\sqrt{n_{S\_E0}} - z_{1-\alpha/2}\sqrt{V^2_{T_S\_H_0}}}{\sqrt{V^2_{T_S\_H_A}}} \quad with \quad 1-\beta = \Phi\left(z_{1-\beta}\right) \quad (2.A.4.2)$$

2.B. Failure probability

The failure probabilities of the Standard and Experimental are respectively indicated by π_{F_St} and π_{F_Ex}, and the sample proportions (p_F) are independent, random binomial variables. The theoretical derivation is very similar and, consequently, has been moved to the paragraph 2.1 of the Appendix 2 of the supplementary material.

3. MODELS FOR THE COMPARISON OF TWO PROPORTIONS

3.A. Success probability (the corresponding treatment of the failure probability is shown in the supplementary material, paragraph 2.2 of the Appendix 2.).

For each Model, we give the null (H_0) and alternative (H_A) hypotheses, the sample distribution, and the formulae for testing H_0 and calculating the sample size and power.

3.A.1. First model (Model 1): Difference between two success probabilities:

3.A.1.1 Null (H_0) and alternative (H_A) hypotheses

The general parameter θ_S becomes $D_{True\_S} = D_{T\_S} = p_{S\_Ex} - p_{S\_St}$ as the difference between the two true probabilities, and δ_{0_S} is the non-inferiority margin (the Greek lower case letter is adopted because it is widely used in the statistical literature).

The non-inferiority hypotheses are:

$$H_0 : p_{S\_Ex} - P_{s\_st} \leq -\delta_{0\_S} \rightarrow D_{T\_S} \leq -\delta_{0\_S}$$

$$H_A : p_{S\_Ex} - P_{s\_st} > -\delta_{0\_S} \rightarrow D_{T\_S} > -\delta_{0\_S}$$

$$\text{where } 0 < \delta_{0\_S} < 1$$

As δ_{0_S} is positive, a negative value ($-\delta_{0\_S}$) is given as it is expected that $\pi_{S\_Ex} < \pi_{S\_S}t$. In addition to these theoretical limits, the upper limit of δ_{0_S} depends on π_{S_St} and is the value of "no clinical/biological difference" compatible with the non-inferiority model. If the lower limit is too near to zero the sample size would be so large that the study would become unfeasible.

The null hypothesis (H_0) is the hypothesis that the difference is less than a negative non-inferiority margin ($-\delta_{0\_S}$), and the alternative hypothesis (H_A) is the hypothesis that the difference is more than the non-inferiority margin, and can consequently be pragmatically considered as clinically or biologically irrelevant.

3.A.1.2 Sampling distribution

It is necessary to consider that, under H_0, the sampling distribution of D_S is shifted: $D_S - \pi_{S\_Ex} - \pi_{S\_St} + \delta_{0\_S}$. Under suitable conditions (generally, large sample sizes) it can be approximated to a Gaussian distribution with mean value and standard error:

$$\mu_{DS} = \pi_{S\_Ex} - \pi_{S\_St} + \delta_{0\_S} \quad and \quad \sigma_{DS} = \sqrt{\frac{\pi_{S\_Ex}\left(1-\pi_{S\_Ex}\right)}{n_{DS\_Ex}} + \frac{\pi_{S\_St}\left(1-\pi_{S\_St}\right)}{n_{DS\_St}}}$$

Its expected values and variances are different under H_0 and H_A, as will be emphasised by using the subscripts $\_H_0$ and $\_H_A$. Under H_0, we have:

$$\mu_{DS\_H_0} = \pi_{S\_Ex\_H_0} - \pi_{S\_St\_H_0} + \delta_{0\_S} = \left(\pi_{S\_St\_H_0} - \delta_{0\_S}\right) - \pi_{S\_St\_H_0} + \delta_{0\_S} = 0$$

and

$$\sigma^2_{DS\_H_0} = \frac{\pi_{S\_Ex\_H_0}\left(1-\pi_{S\_Ex\_H_0}\right)}{n_{DS\_Ex}} + \frac{\pi_{S\_St\_H_0}\left(1-\pi_{S\_St\_H_0}\right)}{n_{DS\_St}}$$

and under the non-inferiority H_A, we have:

$$\mu_{DS\_H_A} = \pi_{S\_Ex\_H_A} - \pi_{S\_St\_H_A} + \delta_{0\_S} = \pi_{S\_Ex} - \pi_{S\_St\_H_A} + \delta_{0\_S} > 0$$

and

$$\sigma^2_{DS\_H_A} = \frac{\pi_{S\_Ex\_H_A}\left(1-\pi_{S\_Ex\_H_A}\right)}{n_{DS\_Ex}} + \frac{\pi_{S\_St}\left(1-\pi_{S\_St\_H_A}\right)}{n_{DS\_St}}$$

As also pointed out by Farrington and Manning [38], it has to be remembered that the unknown probabilities π_{S_Ex} and $\pi_{S\text{-}St}$ in the formulae of the variance under H_0 have to be replaced by the estimates obtained using the observed values (method 1), the values obtained by fixing the marginal totals (method 2), or the maximum likelihood estimates (Method 3).

3.A.1.3 Significance test

The inferiority H_0 is rejected at the $\alpha/2$ significance level if:

$$\frac{D_S - \mu_{DS\_H_0}}{\sigma_{DS\_H_0}} > z_{1-\alpha/2} \rightarrow \frac{p_{S\_Ex} - p_{S\_St} + \delta_{0\_S} - \left(\pi_{S\_Ex\_H_0} - \pi_{S\_St\_H_A} + \delta_{0\_S}\right)}{\sqrt{\frac{\pi_{S\_Ex\_H_0}\left(1-\pi_{S\_Ex\_H_0}\right)}{n_{DS\_Ex}} + \frac{\pi_{S\_St\_H_0}\left(1-\pi_{S\_St\_H_0}\right)}{n_{DS\_St}}}} > z_{1-\alpha/2}$$

or:

$$\frac{p_{S\_Ex} - p_{S\_St} - \left(\pi_{S\_Ex\_H_0} - \pi_{S\_St\_H_A}\right)}{\sqrt{\frac{\pi_{S\_Ex\_H_0}\left(1-\pi_{S\_Ex\_H_0}\right)}{n_{DS\_Ex}} + \frac{\pi_{S\_St\_H_0}\left(1-\pi_{S\_St\_H0}\right)}{n_{DS\_St}}}} > z_{1-\alpha/2}$$

However, a general and easier approach to all of the models is to consider H_0 rejected when the lower limit of the 95% confidence interval is $>-\delta_{S\_0}$ (the non-inferiority margin).

3.A.1.4. Sample size calculation

Let us define the sample sizes in the two treatment groups as $n_{D_{S\_Ex}}$ and $n_{D_{S\_St}}$ with $k = n_{D_{S\_St}} / n_{D_{S\_Ex}}$ and:

$$v^2_{D_S} = n_{D_{S\_Ex}} \cdot \sigma^2_{D_S} = n_{D_{S\_Ex}} \left[\frac{\pi_{S\_Ex}\left(1-\pi_{S\_Ex}\right)}{n_{D_{S\_Ex}}} + \frac{\pi_{S\_St}\left(1-\pi_{S\_St}\right)}{n_{D_{S\_St}}} \right] =$$

$$= \pi_{S\_Ex}\left(1-\pi_{S\_Ex}\right) + \frac{\pi_{S\_St}\left(1-\pi_{S\_St}\right)}{k}$$

which, under H_0 and H_A, become respectively:

$$v^2_{D_{S\_H_0}} = \pi_{S\_Ex\_H_0}\left(1-\pi_{S\_Ex\_H_0}\right) + \frac{\pi_{S\_St\_H_0}\left(1-\pi_{S\_St\_H_0}\right)}{k}$$

$$v^2_{D_{S\_H_A}} = \pi_{S\_Ex\_H_A}\left(1-\pi_{S\_Ex\_H_A}\right) + \frac{\pi_{S\_St\_H_A}\left(1-\pi_{S\_St\_H_A}\right)}{k}$$

Applying formula (2.A.3) at an $\alpha/2$ significance level and (1-β) power gives:

$$n_{DS\_Ex} = \frac{\left(z_{1-\alpha/2}\sqrt{v^2_{DS\_H_0}} + z_{1-\beta}\sqrt{v^2_{DS\_H_A}}\right)^2}{\left(\mu_{DS\_H_A} - \mu_{DS\_H_0}\right)^2} = \frac{\left(z_{1-\alpha/2}\sqrt{v^2_{DS\_H_0}} + z_{1-\beta}\sqrt{v^2_{DS\_H_A}}\right)^2}{\left(\mu_{DS\_H_A}\right)^2} =$$

$$= \frac{\left[z_{1-\alpha/2}\sqrt{\pi_{S\_Ex\_H_0}\left(1-\pi_{S\_Ex\_H_0}\right) + \frac{\pi_{S\_St\_H_0}\left(1-\pi_{S\_St\_H_0}\right)}{k}} + z_{1-\beta}\sqrt{\pi_{S\_Ex\_H_A}\left(1-\pi_{S\_Ex\_H_A}\right) + \frac{\pi_{S\_St\_H_A}\left(1-\pi_{S\_St\_H_A}\right)}{k}}\right]^2}{\left(\pi_{S\_Ex\_H_A} - \pi_{S\_St\_H_A} + \delta_{0\_S}\right)^2} \quad \text{(3.A.1.4)}$$

The denominator of the second term is due to the fact that the expected value of $\mu_{DS\_H_0}$ is zero.

In the case in which $\pi_{S\_Ex\_H_0} = \pi_{S\_Ex\_H_A} = \pi_{S\_Ex}$, and $\pi_{S\_St\_H_0} = \pi_{S\_St\_H_A} = \pi_{S\_St}$, the calculation formula is even simpler:

$$n_{DS\_Ex} = \frac{\left(z_{1-\alpha/2} + z_{1-\beta}\right)^2 \left(\frac{\pi_{S\_St}\left(1-\pi_{S\_St}\right)}{k} + \pi_{S\_Ex}\left(1-\pi_{S\_Ex}\right) \right)}{\left(\pi_{S\_Ex} - \pi_{S\_St} + \delta_{0\_S}\right)^2} \quad \text{(3.A.1.4.Bis)}$$

A further simplification is obtained when k = 1 and $\pi_{S\_Ex} = \pi_{S\_St}$, because the denominator is given by δ_{0_S} squared. The pertinent probability estimates for the sample size calculation have to be entered in the formula.

Formula (3.A.1.4) allows different expected success proportions for the Experimental and Standard, and both Formulae (3.A.1.4) and (3.A.1.4.Bis) allow a different sample size for any k, although k <1 in clinical research as the imbalance is due to randomising more patients to the Experimental in order to obtain more precise estimates.

Once again, it is necessary to use the estimates in accordance with one of the three approaches described by Farrington and Manning [38].

Example 3.A.1.4. (the estimates are calculated using the Farrington and Manning's method 3 [38])

Assuming π_{S_St} = 0.65 and $\pi_{S\_Ex} = \pi_{S\_St}$ = 0.65, and having fixed the non-inferiority margin (θ_{0_S}) at 0.075, with a two-sided significance level (α) of 0.05 leading to $z_{1-\alpha/2}$= 1.96, and a power (1-β) of 0.90 leading to a $z_{1-\beta}$ = 1.2816 with $(z_{1-\alpha/2}+z_{1-\beta})^2$ = 10.507971, the sample size for the experimental group, which is equal to that of the standard group (k = 1) is easily obtained from Formula 3.A.1.4 as n_{DS_Ex} = 849.98 ≈ 850.

It is worth repeating that, in non-inferiority settings, a very sensible starting point is to put $\pi_{S\_St} = \pi_{S\_Ex}$, thus assuming that the Experimental is at least as effective as the Standard, because this is in line with the equipoise position that allows ethically feasible randomised controlled trials. If the effectiveness of the two drugs is considered to be different, adding the non-inferiority margin (δ_{0_S}) could lead to a clinically relevant total difference that is unsuitable for a non-inferiority study.

3.A.1.5. Power calculation

Under the alternative hypothesis that $D_{T\_S} > -\delta_{0\_S}$, the power of the above test is:

$$1-\beta=\Phi\left(\frac{\left(\mu_{D_{S\_H_A}}-\mu_{D_{S\_H_0}}\right)-z_{1-\alpha/2}\sigma_{D_{S\_H_0}}}{\sigma_{D_{S\_H_A}}}\right)=\Phi\left(\frac{\left(\mu_{D_{S\_H_A}}-0\right)-z_{1-\alpha/2}\sigma_{D_{S\_H_0}}}{\sigma_{D_{S\_H_A}}}\right)=$$

$$=\Phi\left(\frac{\left(\pi_{S\_Ex\_H_A}-\pi_{S\_St\_H_A}+\delta_{0\_S}\right)-z_{1-\alpha/2}\sqrt{\frac{\pi_{S\_Ex\_H_0}\left(1-\pi_{S\_Ex\_H_0}\right)}{n_{D_{S\_Ex}}}+\frac{\pi_{S\_St\_H_0}\left(1-\pi_{S\_St\_H_0}\right)}{n_{D_{S\_St}}}}}{\sqrt{\frac{\pi_{S\_Ex\_H_A}\left(1-\pi_{S\_Ex\_H_A}\right)}{n_{D_{S\_Ex}}}+\frac{\pi_{S\_St\_H_A}\left(1-\pi_{S\_St\_H_A}\right)}{n_{D_{S\_St}}}}}\right) \quad (3.A.1.5)$$

Alternatively, solving $z_{1-\beta}$ and then calculating the probability value of this quantile, we have:

$$z_{1-\beta}=\frac{\left(\pi_{S\_Ex}-\pi_{S\_St\_H_A}+\delta_{0\_S}\right)\sqrt{n_{S\_Ex}}-z_{1-\alpha/2}\sqrt{\pi_{S\_Ex\_H_0}\left(1-\pi_{S\_Ex\_H_0}\right)+\frac{\pi_{S\_St\_H_0}\left(1-\pi_{S\_St\_H_0}\right)}{k}}}{\sqrt{\pi_{S\_Ex\_H_A}\left(1-\pi_{S\_Ex\_H_A}\right)+\frac{\pi_{S\_St\_H_A}\left(1-\pi_{S\_St\_H_A}\right)}{k}}} \quad (3.A.1.5.Bis)$$

$$then,\ 1-\beta=\Phi\left(z_{1-\beta}\right)$$

3.A.2. Second Model: ratio between two success probabilities (the term "relative risk" should be restricted to the ratio of two failure probabilities).

3.A.2.1.1- Null (H_{0_S}) and alternative (H_{A_S}) hypotheses

The general parameter θ_S becomes $R_{T\_S}=\frac{\pi_{S\_Ex}}{\pi_{S\_St}}$, with R_{0_S} the non-inferiority limit of R_{T_S}, and the null and alternative hypotheses are:

$$H_0: R_{T\_S} \le R_{0\_S} \rightarrow \pi_{S\_Ex} - R_{0\_S}\pi_{S\_St} \le 0$$

$$H_A: R_{T\_S} > R_{0\_S} \rightarrow \pi_{S\_Ex} - R_{0\_S}\pi_{S\_St} > 0$$

As it is expected that $\pi_{S\_Ex} < \pi_{S\_St}$, we have $0 < R_{0\_S} < 1$.

The null hypothesis (H_0) is the hypothesis of inferiority given by a ratio that is less than the non-inferiority margin R_{0_S}, and the alternative hypothesis (H_A) is the hypothesis of a ratio that is more than the non-inferiority margin and, consequently, a clinically or biologically irrelevant value.

3.A.2.1.2. Sampling distribution

It is necessary to consider the sampling distribution of $R_{T\_S}=\frac{p_{S\_Ex}}{p_{S\_St}}$ and, in accordance with the second formulation of the above H_0 and H_A, we consider a sampling distribution of $R_S = p_{S\_Ex} - R_{0\_S}p_{S\_St}$ with an expected value $E(_{RS})$ and standard deviation (σ_{RS}) respectively given by:

$$E(R_S)=\mu_{R_S}=\pi_{S\_Ex}-R_{0\_S}\pi_{S\_St} \quad and \quad \sigma_{R_S}=\left(\frac{\pi_{S\_Ex}\left(1-\pi_{S\_Ex}\right)}{n_{R_{S\_Ex}}}+R^2{}_{0\_S}\frac{\pi_{S\_St}\left(1-\pi_{S\_St}\right)}{n_{R_{S\_St}}}\right)^{1/2}$$

Under suitable conditions (generally, large sample sizes), the statistical test can be referred to a standardised Gaussian distribution [Z(0; 1)]:

$$\frac{R_S-\mu_{R_S}}{\sigma_{R_S}}\approx Z(0,1)\rightarrow\frac{p_{S\_Ex}-R_{0\_S}p_{S\_St}-\left(\pi_{S\_Ex}-R_{0\_S}\pi_{S\_St}\right)}{\sqrt{\frac{\pi_{S\_Ex}\left(1-\pi_{S\_Ex}\right)}{n_{R_{S\_Ex}}}+R^2_{0\_S}\frac{\pi_{S\_St}\left(1-\pi_{S\_St}\right)}{n_{R_{S\_St}}}}}\approx Z(0,1)$$

3.A.2.1.3. Significance test

H_0 at an $\alpha/2$ level of significance is rejected if:

$$\frac{R_S-\mu_{R_{S\_H_0}}}{\sigma_{R_{S\_H_0}}}>z_{1-\alpha/2}$$

$$\rightarrow\frac{p_{S\_Ex}-R_{0\_S}p_{S\_St}-\left(\pi_{S\_Ex\_H_0}-R_{0\_S}\pi_{S\_St\_H_0}\right)}{\sqrt{\frac{\pi_{S\_Ex\_H_0}\left(1-\pi_{S\_Ex\_H_0}\right)}{n_{R_{S\_Ex}}}+R^2_{0\_S}\frac{\pi_{S\_St\_H_0}\left(1-\pi_{S\_St\_H_0}\right)}{n_{R_{S\_St}}}}}>z_{1-\alpha/2}\qquad(3.A.2.1.3)$$

3.A.2.1.4. Sample size calculation

Let us indicate the sample sizes of the two treatment groups as n_{RS_Ex} and n_{RS_St} with $k=n_{R_{S\_St}}/n_{R_{S\_Ex}}$ and:

$$v^2_{R_S}=n_{R_{S\_Ex}}\cdot\sigma^2_{R_S}=n_{R_{S\_Ex}}\left[\frac{\pi_{S\_Ex}\left(1-\pi_{S\_Ex}\right)}{n_{R_{S\_Ex}}}+R^2_{0\_S}\frac{\pi_{S\_St}\left(1-\pi_{S\_St}\right)}{n_{R_{S\_St}}}\right]=$$

$$=\pi_{S\_Ex}\left(1-\pi_{S\_Ex}\right)+R^2_{0\_S}\frac{\pi_{S\_St}\left(1-\pi_{S\_St}\right)}{k}$$

which, under H_0 and H_A, respectively become:

$$v^2_{R_{S\_H_0}}=\pi_{S\_Ex\_H_0}\left(1-\pi_{S\_Ex\_H_0}\right)+R^2_{0\_S}\frac{\pi_{S\_St\_H_0}\left(1-\pi_{S\_St\_H_0}\right)}{k}$$

$$v^2_{R_{S\_H_A}}=\pi_{S\_Ex\_H_A}\left(1-\pi_{S\_Ex\_H_A}\right)+R^2_{0\_S}\frac{\pi_{S\_St\_H_A}\left(1-\pi_{S\_St\_H_A}\right)}{k}$$

Using Formula (2.A.3), we have

$$n_{R_{S\_Ex}}=\frac{\left(z_{1-\alpha/2}\sqrt{v^2_{R_{S\_H_0}}}+z_{1-\beta}\sqrt{v^2_{R_{S\_H_A}}}\right)^2}{\left(\mu_{R_{S\_H_A}}-\mu_{R_{S\_H_0}}\right)^2}=$$

$$=\frac{\left(z_{1-\alpha/2}\sqrt{\pi_{S\_Ex\_H_0}\left(1-\pi_{S\_Ex\_H_0}\right)+R^2_{0\_S}\frac{\pi_{S\_St\_H_0}\left(1-\pi_{S\_St\_H_0}\right)}{k}}+z_{1-\beta}\sqrt{\pi_{S\_Ex\_H_A}\left(1-\pi_{S\_Ex\_H_A}\right)+R^2_{0\_S}\frac{\pi_{S\_St\_H_A}\left(1-\pi_{S\_St\_H_A}\right)}{k}}\right)^2}{\left(\pi_{S\_Ex\_H_A}-R_{0\_S}\cdot\pi_{S\_St\_H_A}\right)^2}\qquad(3.A.2.1.4)$$

In the case of $\pi_{S\_Ex\_H_0}=\pi_{S\_Ex\_H_A}=\pi_{S\_Ex}$ and $\pi_{S\_St\_H_0}=\pi_{S\_St\_H_A}=\pi_{S\_St}$, Formula 3.A.2.1.4 becomes:

$$n_{R_{S\_Ex}}=\frac{\left(z_{1-\alpha}+z_{1-\beta}\right)^2\left(\pi_{S\_Ex}\left(1-\pi_{S\_Ex}\right)+R^2_{0\_S}\frac{\pi_{S\_St}\left(1-\pi_{S\_St}\right)}{k}\right)}{\left(\pi_{S\_Ex}-R_{0\_S}\pi_{S\_St}\right)^2}\qquad(3.A.2.1.4.Bis)$$

$$where\ k=n_{R_{S\_St}}/n_{R_{S\_Ex}}\rightarrow n_{R_{S\_St}}=k\cdot n_{R_{S\_Ex}}$$

Example 3.A.2.1.4 (the estimates are calculated using Farrington and Manning's method 3 [38])

Assuming $\pi_{S\_St} = 0.65$ and $\pi_{S\_Ex} = \pi_{S\_St} = 0.65$ (or $R_{T\_S}=1$), and having fixed the non-inferiority margin as $R_{0\_S} = 0.8846154$, the two-sided significance level (α) at 0.05 leading $z_{1-\alpha/2} = 1.96$, and power $(1 - \beta) = 0.90$ leading to a $z_{1-\beta} = 1.2816$ with $(z_{1-\alpha/2}+z_{1-\beta})^2 = 10.507971$, the sample size for the Experimental, which is equal to that of the Standard ($k = 1$) is easily obtained from Formula 3.A.2.4 as $n_{RS\_Ex} = 757.52 \approx 758$.

It should be noted that the above conditions correspond to those of Example 3.A.1.4 for model 1 as R_{0_S} is obtained from δ_{0_S} by means of the conversion formula: $R_{0\_S} = (\pi_{S\_St} - \delta_{0\_S}) / \pi_{S\_St}$. In this case, the sample size of 758 in each treatment group is much lower than the 850 obtained using model 1.

Finally, it should be pointed out that, if R_{0_S} is rounded to 0.885, $n_{RS\_Ex} = 762.89 \approx 763$: i.e. an increase of just four-tenths of a thousand in the non-inferiority margin expressed as the ratio leads to an increase of five subjects in each treatment group.

3.A.2.1.5. Power calculation

Under the alternative hypothesis that $R_{T\_S} > R_{0\_S}$, the power of the above test is:

$$1-\beta = \Phi\left(\frac{\left(\mu_{RS\_H_A} - \mu_{RS\_H_0}\right) - z_{1-\alpha}\sigma_{RS\_H_0}}{\sigma_{RS\_H_A}}\right) =$$

$$= \Phi\left(\frac{\left(\pi_{S\_Ex\_H_A} - R_{0\_S}\pi_{S\_St\_H_A}\right) - z_{1-\alpha}\sqrt{\frac{\left(R_{0\_S}\pi_{S\_St\_H_0}\right)\left(1-R_{0\_S}\pi_{S\_St\_H_0}\right)}{n_{RS\_Ex}} + \frac{\pi_{S\_St\_H_0}\left(1-\pi_{S\_St\_H_0}\right)}{n_{RS\_St}}}}{\sqrt{\frac{\pi_{S\_Ex\_H_A}\left(1-\pi_{S\_Ex\_H_A}\right)}{n_{RS\_Ex}} + \frac{\pi_{S\_St\_H_A}\left(1-\pi_{S\_St\_H_A}\right)}{n_{RS\_St}}}}\right) \quad (3.A.2.1.5)$$

Alternatively, by solving $z_{1-\beta}$ and then calculating the probability value associated with this quantile, we have:

$$z_{1-\beta} = \frac{\left(\pi_{S\_Ex\_H_A} - R_{0\_S}\cdot\pi_{S\_St\_H_A}\right)\sqrt{nR_{S\_Ex}} - z_{1-\alpha/2}\sqrt{\pi_{S\_Ex\_H_0}\left(1-\pi_{S\_Ex\_H_0}\right) + R^2_{0\_S}\frac{\pi_{S\_St\_H_0}\left(1-\pi_{S\_St\_H_0}\right)}{k}}}{\sqrt{\pi_{S\_Ex\_H_A}\left(1-\pi_{S\_Ex\_H_A}\right) + R^2_{0\_S}\frac{\pi_{S\_St\_H_A}\left(1-\pi_{S\_St\_H_A}\right)}{k}}} \quad (3.A.2.1.5.Bis)$$

$$1-\beta = \Phi^{-1}(z_{1-\beta})$$

3.A.2.2.1. Second Model Extension: (natural) logarithm of the ratio between two success probabilities.

With $LR_{T\_S}=\ln(R_{T\_S})$ and $LR_{0\_S}=\ln(R_{0\_S})$, the previous null and alternative hypotheses concerning the ratio between two success probabilities are re-written with the log-transformation as:

$$H_0 : LR_{T\_S} \leq LR_{0\_S}$$
$$H_A : LR_{T\_S} > LR_{0\_S}$$

As $0 < R_{0\_S} < 1$, we have $LR_{0\_S} < 0$.

3.A.2.2.2 Sampling distribution

It is necessary to consider the shifted sampling distribution of $LR_S=\ln(P_{S\_EX})-\ln(P_{S\_ST})-LR_{0\_S}$ The following formulae show the expected value $E(LR_S)$, with the standard deviation (σ_{LRS}) being calculated using the delta method:

$$E(LR_S) = \mu_{LR_S} \approx \ln(\pi_{S\_Ex}) - \ln(\pi_{S\_St}) - LR_{0\_S} \quad and \quad \sigma_{LR_S} \approx \left(\frac{1-\pi_{S\_Ex}}{n_{LR_{S\_Ex}} \pi_{S\_Ex}} + \frac{1-\pi_{S\_St}}{n_{LR_{LR\_St}} \pi_{S\_St}} \right)^{1/2}$$

For large sample sizes, this distribution can be approximated to a standardised Gaussian formula:

$$\frac{LR_S - E(LR_S)}{\sigma_{LR_S}} \approx Z(0;1) \rightarrow \frac{LR_S - \left[\ln(\pi_{S\_Ex}) - \ln(\pi_{S\_St}) - LR_{0\_S} \right]}{\sqrt{\frac{1-\pi_{S\_Ex}}{n_{LR_{S\_Ex}} \pi_{S\_Ex}} + \frac{1-\pi_{S\_St}}{n_{LR_{S\_St}} \pi_{S\_St}}}} \approx Z(0,1)$$

3.A.2.2.3 Significance test

The null hypothesis is rejected at a $\alpha/2$ level of significance if:

$$\frac{LR_S - \mu_{LR_{S\_H_0}}}{\sigma_{LR_{S\_H_0}}} > z_{1-\alpha/2} \rightarrow \frac{LR_S - \left[\ln(R_{T\_S}) - LR_{0\_S} \right]}{\sqrt{\frac{1-\pi_{S\_Ex\_H_0}}{n_{LR_{S\_Ex}} \pi_{S\_Ex\_H_0}} + \frac{1-\pi_{S\_St\_H_0}}{n_{LR_{S\_St}} \pi_{S\_St\_H_0}}}} > z_{1-\alpha/2}$$

3.A.2.2.4 Sample size calculation

Let the sample sizes of the two treatment groups be n_{LRS_Ex} and n_{LRS_St} with $k = n_{LR_{S\_St}} / n_{LR_{S\_Ex}}$ and:

$$v^2_{LR_S} = n_{LR_{S\_Ex}} \cdot \sigma^2_{LR_S} = n_{LR_{S\_Ex}} \left(\frac{1-\pi_{S\_Ex}}{n_{LR_{S\_Ex}} \cdot \pi_{S\_Ex}} + \frac{1-\pi_{S\_St}}{n_{LR_{S\_St}} \cdot \pi_{S\_St}} \right) = \frac{1-\pi_{S\_Ex}}{\pi_{S\_Ex}} + \frac{1-\pi_{S\_St}}{k \cdot \pi_{S\_St}}$$

which, under H_0 and H_A respectively become:

$$v^2_{LR_{S\_H_0}} = \frac{1-\pi_{S\_Ex\_H_0}}{\pi_{S\_Ex\_H_0}} + \frac{1-\pi_{S\_St\_H_0}}{k \cdot \pi_{S\_St\_H_0}}$$

$$v^2_{LR_{S\_H_A}} = \frac{1-\pi_{S\_Ex\_H_A}}{\pi_{S\_Ex\_H_A}} + \frac{1-\pi_{S\_St\_H_A}}{k \cdot \pi_{S\_St\_H_A}}$$

From the general formula 2.A.3, we obtain:

$$n_{LR_{S\_Ex}} = \frac{\left(z_{1-\alpha/2} \sqrt{v^2_{LR_{S\_H_0}}} + z_{1-\beta} \sqrt{v^2_{LR_{S\_H_A}}} \right)^2}{\left(\mu_{LR_{S\_H_A}} - \mu_{LR_{S\_H_0}} \right)^2} =$$

$$= \frac{\left(z_{1-\alpha/2} \sqrt{\frac{1-\pi_{S\_Ex\_H_0}}{\pi_{S\_Ex\_H_0}} + \frac{1-\pi_{S\_St\_H_0}}{k \cdot \pi_{S\_St\_H_0}}} + z_{1-\beta} \sqrt{\frac{1-\pi_{S\_Ex\_H_A}}{\pi_{S\_Ex\_H_A}} + \frac{1-\pi_{S\_St\_H_A}}{k \cdot \pi_{S\_St\_H_A}}} \right)^2}{\left[\ln(\pi_{S\_Ex\_H_A}) - \ln(\pi_{S\_St\_H_A}) - LR_{0\_S} \right]^2} \qquad 3.A.2.2.4$$

Again, in the case of $\pi_{S\_Ex\_H_0} = \pi_{S\_Ex\_H_A} = \pi_{S\_Ex}$, and $\pi_{S\_St\_H_0} = \pi_{S\_St\_H_A} = \pi_{S\_St}$ it is straightforward to obtain the simpler sample size calculation formula.

Example 3.A.2.2.4 (the estimates are obtained using Farrington and Manning's method 3 [38].)

Assuming $\pi_{S\_St} = 0.65$ and $\pi_{S\_Ex} = \pi_{S\_St} = 0.65$ (or $R_{T\_S} = 1$), and having fixed the non-inferiority margin $L_{RO\_S} = -0.1226$, a significance level (α) at 0.05 two-sided leading to $z_{1-\alpha/2} = 1.96$ and a power of 0.90 leading to a $z_{1-\beta} = 1.2816$ with $(z_{1-\alpha/2} + z_{1-\beta})^2 = 10.507971$, the sample size for the Experimental group, which is equal to that of the Standard group (k = 1), is easily obtained from the formula 3.A.2.1.2 as $n_{LRS\_Ex} = 752.80 \approx 753$.

It has to be noted that the above conditions correspond to those of Example 3.A.1.4 of model 1 insofar as R_{O_S} is obtained δ_{O_S} by means of the pertinent conversion formula: $R_{O\_S} = (\pi_{S\_St} - \delta_{O\_S})/\pi_{S\_St}$ and $L_{RO\_S} = \ln(R_{O\_S}) = \ln(0.884645) = -0.1226$. In this case, the sample size of 753 in each treatment group is a little lower than the 758 required by Model 2, and much lower than the 850 required by model 1.

3.A.2.2.5 Power Calculation

Under the alternative hypothesis that $L_{RT\_S} > L_{RO\_S}$, the power of the above test is:

$$1-\beta = \Phi\left(\frac{\left(\mu_{LRS\_H_A} - \mu_{LRS\_H_0}\right) - z_{1-\alpha}\sigma_{LRS\_H_0}}{\sigma_{LRS\_H_A}}\right) = \Phi\left(\frac{\left(\pi_{S\_Ex\_H_A} - R_{O\_S}\pi_{S\_St}\right) - z_{1-\alpha}\sqrt{\frac{1-\pi_{S\_Ex}}{n_{LRS\_Ex}\pi_{S\_Ex}} + \frac{1-\pi_{S\_St}}{n_{LRS\_St}\pi_{S\_St}}}}{\sqrt{\frac{1-\pi_{S\_Ex\_H_A}}{n_{LRS\_Ex}\pi_{S\_Ex\_H_A}} + \frac{1-\pi_{S\_St}}{n_{LRS\_St}\pi_{S\_St}}}}\right) \quad (3.A.2.2.5)$$

Alternatively, Formula 3.A.2.1.4 can be straightforwardly solved using $z_{1-\beta}$, after which the corresponding probability value can be calculated.

3.A.3. Third Model: Odds ratio (OR_S) of two success probabilities and its (natural) logarithm (LOR_S).

The comparison of two success proportions can also be expressed using the odds ratio (OR_S), with the true odds ratio (OR_{T_S}):

$$(OR_{T\_S}):\ OR_{T\_S} = \frac{\pi_{S\_Ex}}{1-\pi_{S\_Ex}} \Big/ \frac{\pi_{S\_St}}{1-\pi_{S\_St}}$$

and OR_{O_S} as its non-inferiority margin.

3.A.3.1 Null (H_0) and alternative (H_A) hypotheses

The non-inferiority hypotheses in terms of the odds ratio are:

$$H_0 : OR_{T\_S} \le OR_{0\_S}$$
$$H_A : OR_{T\_S} > OR_{0\_S}$$

with $0 < OR_{O\_S} < 1$ as it is expected that $\pi_{S\_Ex} < \pi_{S\_St}$.

Of course, the actual limits are the values compatible with the non-inferiority settings and, consequently, the "clinically or biologically irrelevant difference".

The sample odds ratio (OR_S) is given by:

$$OR_S = \frac{p_{S\_Ex}}{1-p_{S\_Ex}} \Big/ \frac{p_{S\_St}}{1-p_{S\_St}} = \frac{p_{S\_Ex}}{q_{S\_Ex}} \Big/ \frac{p_{S\_St}}{q_{S\_St}}$$

However, it is better to consider its natural logarithm, given by $LOR_S = \ln(OR_S)$ because of its more suitable distributional properties: consequently, we have $LOR_{T\_S} = \ln(OR_{T\_S})$ and $LOR_{O\_S} = \ln(OR_{O\_S})$ with the hypotheses:

$$H_0 : LOR_{T\_S} \leq LOR_{0\_S}$$
$$H_A : LOR_{T\_S} > LOR_{0\_S}$$

As $0 < OR_{0\_S} < 1$, LOR_{0_S} is < 0.

3.A.3.2 Sampling distribution

It is necessary to consider the shifted sampling distribution of:

$$LOR_S = ln(OR_S) - LOR_{0\_S} = ln(p_{S\_Ex}) - ln(q_{S\_Ex}) - ln(p_{S\_St}) + ln(q_{S\_St}) - LOR_{0\_S}$$
$$where: q_{S\_Ex} = 1 - p_{S\_Ex} \ and \ q_{S\_St} = 1 - p_{S\_St}$$

The following formulae show the expected value E(LOR_S), with the standard deviation (σ_{LORS}) being calculated using the delta method:

$$E(LOR_S) = \mu_{LOR_S} \approx ln(\pi_{S\_Ex}) - ln(1-\pi_{S\_Ex}) - ln(\pi_{S\_St}) + ln(1-\pi_{S\_St}) - LOR_{0\_S}$$
$$\sigma_{LOR_S} \approx \left(\frac{1}{nLOR_{S\_Ex}\pi_{S\_Ex}(1-\pi_{S\_Ex})} + \frac{1}{nLOR_{S\_St}\pi_{S\_St}(1-\pi_{S\_St})} \right)^{1/2}$$

By expressing π_{S_Ex} from the known parameters of the model (OR_{T_S} and π_{S_St}), we obtain:

$$\pi_{S\_Ex} = \frac{OR_{T\_S}\pi_{S\_St}}{1 + \pi_{S\_St}(OR_{T\_S} - 1)}$$

Assuming that the distribution of the LORs for a large sample size can be approximated using a standardised Gaussian curve [Z(0, 1)], we have:

$$\frac{LOR_S - E(LOR_S)}{\sigma_{LOR_S}} \approx Z(0,1) \rightarrow$$
$$\frac{LOR_S - \left[ln(\pi_{S\_Ex}) - ln(1-\pi_{S\_Ex}) - ln(\pi_{S\_St}) + ln(1-\pi_{S\_St}) - LOR_{0\_S} \right]}{\sqrt{\frac{1}{nLOR_{S\_Ex}\pi_{S\_Ex}(1-\pi_{S\_Ex})} + \frac{1}{nLOR_{S\_St}\pi_{S\_St}(1-\pi_{S\_St})}}} \approx Z(0,1)$$

3.A.3.3 Statistical test

The null hypothesis is rejected at the $\alpha/2$ level of significance if:

$$\frac{LOR_S - \mu_{LOR_{S\_H_0}}}{\sigma_{LOR_{S\_H_0}}} > z_{1-\alpha/2} \rightarrow \frac{LOR_S - \left[ln(\pi_{S\_Ex\_H_0}) - ln(1-\pi_{S\_Ex\_H_0}) - ln(\pi_{S\_St\_H_0}) + ln(1-\pi_{S\_St\_H_0}) - LOR_{0\_S} \right]}{\sqrt{\frac{1}{nLOR_{S\_Ex}\pi_{S\_Ex\_H_0}(1-\pi_{S\_Ex\_H_0})} + \frac{1}{nLOR_{S\_St}\pi_{S\_St\_H_0}(1-\pi_{S\_St\_H_0})}}} > z_{1-\alpha/2}$$

3.A.3.4 Sample size calculation

Let the sample sizes in the two treatment groups be $n_{LOR_{S\_Ex}}$ and $n_{LOR_{S\_St}}$ with $k = n_{LOR_{S\_St}} / n_{LOR_{S\_Ex}}$, and:

$$v^2_{LOR_S} = n_{LOR_{S\_Ex}} \cdot \sigma^2_{LOR_S} = n_{LOR_{S\_Ex}} \left(\frac{1}{n_{LOR_{S\_Ex}} \cdot \pi_{S\_Ex}\left(1-\pi_{S\_Ex}\right)} + \frac{1}{n_{LOR_{S\_St}} \cdot \pi_{S\_St}\left(1-\pi_{S\_St}\right)} \right) =$$

$$= \frac{1}{\pi_{S\_Ex}\left(1-\pi_{S\_Ex}\right)} + \frac{1}{k \cdot \pi_{S\_St}\left(1-\pi_{S\_St}\right)}$$

which respectively become under H_0 and H_A:

$$v^2_{LOR_{S\_H_0}} = \frac{1}{\pi_{S\_Ex\_H_0}\left(1-\pi_{S\_Ex\_H_0}\right)} + \frac{1}{k \cdot \pi_{S\_St\_H_0}\left(1-\pi_{S\_St\_H_0}\right)}$$

$$v^2_{LOR_{S\_H_A}} = \frac{1}{\pi_{S\_Ex\_H_A}\left(1-\pi_{S\_Ex\_H_A}\right)} + \frac{1}{k \cdot \pi_{S\_St\_H_A}\left(1-\pi_{S\_St\_H_A}\right)}$$

From the general formula (2.A.3), we obtain:

$$n_{LOR_{S\_Ex}} = \frac{\left(z_{1-\alpha}\sqrt{v^2_{LOR_{S\_H_0}}} + z_{1-\beta}\sqrt{v^2_{LOR_{S\_H_A}}}\right)^2}{\left(\mu_{LOR_{S\_H_A}} - \mu_{LOR_{S\_H_0}}\right)^2} =$$

$$= \frac{\left(z_{1-\alpha}\sqrt{\frac{1}{\pi_{S\_Ex\_H_0}\left(1-\pi_{S\_Ex\_H_0}\right)} + \frac{1}{k \cdot \pi_{S\_St\_H_0}\left(1-\pi_{S\_St\_H_0}\right)}} + z_{1-\beta}\sqrt{\frac{1}{\pi_{S\_Ex\_H_A}\left(1-\pi_{S\_Ex\_H_A}\right)} + \frac{1}{k \cdot \pi_{S\_St\_H_A}\left(1-\pi_{S\_St\_H_A}\right)}}\right)^2}{\left[\ln\left(\pi_{S\_Ex\_H_A}\right) - \ln\left(1-\pi_{S\_Ex\_H_A}\right) - \ln\left(\pi_{S\_St\_H_A}\right) + \ln\left(1-\pi_{S\_St\_H_A}\right) - LOR_{0\_S}\right]^2} \quad \text{3.A.3.4}$$

Again, if $\pi_{S\_Ex\_H_0} = \pi_{S\_Ex\_H_A} = \pi_{S\_Ex}$ and $\pi_{S\_St\_H_0} = \pi_{S\_St\_H_A} = \pi_{S\_St}$, it is easy to obtain a simpler formula.

Example 3.A.3.4 (the estimates are calculated using Farrington & Manning's method 3 [38].)

With π_{S_St} = 0.65 and π_{S_Ex} = 0.65, a fixed non-inferiority margin OR_{0_S} of 0.728507 (from [0.575.(1 - 0.65)] / [0.65.(1 - 0.575)], being 0.575 = 0.65 – 0.075) or LOR_{0_S} = ln(LOR_{0_S}) = ln(0.728507) = -0.316758, a two-sided significance level (α) of 0.05 leading to $z_{1-\alpha/2}$ = 1.96 and a power of 0.90 leading to a $z_{1-\beta}$ = 1.2816 with $(z_{1-\alpha/2} + z_{1-\beta})^2$ = 10.507971, the sample size for the experimental group, which is equal to that of the standard group (k = 1), is easily obtained from the formula (3.A.3.3) as $nLOR_{S\_Ex}$ = 920.64 ≈ 921. This is larger than the 850 calculated using the difference (Model 1), and much larger than the 758 calculated using the ratio (R_S, model 2) or the 753 obtained from the LR_S (model 2.1).

3.A.3.5 Power calculation

Under the alternative hypothesis that $OR_{T\_S} > OR_{0\_S}$, the power of the above test is:

$$1-\beta = \Phi\left(\frac{\left(\mu_{LOR_{S\_H_A}} - \mu_{LOR_{S\_H_0}}\right) - z_{1-\alpha/2}\sigma_{LOR_{S\_H_0}}}{\sigma_{LOR_{S\_H_A}}}\right) =$$

$$= \Phi\left(\frac{\left[\ln\left(\pi_{S\_Ex\_H_A}\right) - \ln\left(1-\pi_{S\_Ex\_H_A}\right) - \ln\left(\pi_{S\_St\_H_A}\right) + \ln\left(1-\pi_{S\_St\_H_A}\right) - LOR_{0\_S}\right] - z_{1-\alpha/2}\sqrt{\frac{1}{n_{LOR_{S\_Ex}}\pi_{S\_Ex\_H_0}\left(1-\pi_{S\_Ex\_H_0}\right)} + \frac{1}{n_{LOR_{S\_St}}\pi_{S\_St\_H_0}\left(1-\pi_{S\_St\_H_0}\right)}}}{\sqrt{\frac{1}{n_{LOR_{S\_Ex}}\pi_{S\_Ex\_H_A}\left(1-\pi_{S\_Ex\_H_A}\right)} + \frac{1}{n_{LOR_{S\_St}}\pi_{S\_St\_H_A}\left(1-\pi_{S\_St\_H_A}\right)}}}\right) \quad (3.A.3.5)$$

Otherwise (and equivalently), Formula 3.A.3.4 can be straightforwardly solved using $z_{1-\beta}$, after which the probability value associated with this quantile can be calculated.

3.A.4. Success probability tables

Table 1 shows the H_{0_S} hypothesis, sample distributions and sample size calculation formulae of the three models, with the second model being divided into the ratio (R_S, model 2.1) and logarithm of the ratio models (LR_S, model 2.2), and OR_S and LOR_S being considered together as the third model.

Table 1.1 shows the sample sizes calculated, using the three methods of Farrington and Manning [38] for R_{T_S} values ranging from 0.85 to 1.0 at intervals of 0.05 and π_{S_St} values ranging from 0.1 to 0.9 at intervals of 0.2, assuming that $\alpha = 0.025$, $1\text{-}\beta = 0.80$, and the non-inferiority margins (expressed as R_{0_S} in order to be consistent with Laster et al.) [49] are 0.8, 0.85 and 0.95.

For example, with $\pi_{S\_St} = 0.5$ and $\pi_{S\_EX} = 0.425$ (giving $R_{T\_S} = 0.85$, $\delta_{0\_S} = 0.10$, $\alpha = 0.025$ and $1\text{-}\beta = 0.80$), it is first necessary to calculate $R_{0\_S} = 1 - \delta_{0\_S} / \pi_{S\_St} = 1 - 0.10/0.5 = 0.80$, and then it is possible to read in the row with π_{S_St} 0.5 and R_{0_S} 0.8 and M = 1 (for the method) that $n_{DS\_Ex} = 6{,}208$, $n_{RS\_Ex} = 5{,}078$, $n_{LRS\_Ex} = 5{,}025$, and $n_{LORS\_Ex} = 5{,}965$ in columns 4-8 (corresponding to $R_{T\_S} = 0.85$). The subsequent two rows show the sample sizes calculated using methods 2 and 3. It is important to note that, when $\pi_{S\_St} = \pi_{S\_EX} = 0.5$, giving $R_{T\_S} = 1.00$, the sample sizes become 392, 322, 315, and 382.

4. SWITCHING NON-INFERIORITY MARGINS FROM ONE MODEL TO ANOTHER: COMPARISON OF SAMPLE SIZES

We propose a general method that allows to switch from one model to another, valid for all four models for both successes and failures. As a consequence, the hypotheses of a model are re-parametrized in those of another model, obtaining the corresponding non-inferiority margins. To maintain consistency in definitions and approaches it is necessary to place the general constraint that the NI margin of the final model is calculated only by the NI margin of the starting model and by π_{S_St} which is independent of the NI margin and considered "known" in the planning phase of a study.

4.A. Success probability

Switching the calculation of non-inferiority margins from one model to another is based on a fixed π_{S_St}, which is considered as a "known" during the planning phase of a trial in the same way as the true difference (D_{T_S}), the true ratio (R_{T_S}) and the true odds ratio (OR_{T_S}).

Table 3 shows the switching formulae of the three models.

4.A.1. Switching non-inferiority margins from one model to another

4.A.1.1 Model 1 (difference/delta: D_S) vs model 2.1 (ratio: R_S) and model 2.2 (ln ratio: LR_S)

Starting from the non-inferiority H_{0_S} hypothesis concerning the difference between two success probabilities ($H_{0\_S}: \pi_{S\_Ex} - \pi_{S\_St} \leq -\delta_{0\_S}$), the last term of the last inequality is obtained by dividing both terms of the first inequality by π_{S_St}:

$$\pi_{S\_Ex} - \pi_{S\_St} \leq -\delta_{0\_S} \rightarrow \frac{\pi_{S\_Ex}}{\pi_{S\_St}} - 1 \leq -\frac{\delta_{0\_S}}{\pi_{S\_St}} \rightarrow \frac{\pi_{S\_Ex}}{\pi_{S\_St}} \leq 1 - \frac{\delta_{0\_S}}{\pi_{S\_St}}$$

Thus, by putting $R_{0\_S} = 1 - \frac{\delta_{0\_S}}{\pi_{S\_St}}$, it is possible to calculate the non-inferiority margin for the ratio of two success probabilities.

Then, using $R_{T\_S} = \frac{\pi_{S\_Ex}}{\pi_{S\_St}}$ and $R_{0\_S} = 1 - \frac{\delta_{0\_S}}{\pi_{S\_St}}$, the statistical hypotheses can be formulated in the terms of a ratio. In the case of H_{0_S}, it can be written:

$$H_{0\_S}: \pi_{S\_Ex} - \pi_{S\_St} \leq -\delta_{0\_S} \rightarrow H_0: R_{T\_S} \leq R_{0\_S}$$

Alternatively, starting from the non-inferiority hypotheses of the ratio of two success probabilities (R_{0_S}), it is necessary to fix the value of π_{S_St} in order to obtain the non-inferiority margin in terms of a difference: $_{\delta 0\_S} = \pi_{S\_St} (1 - R_{0\_S})$. Obviously, $\delta_{0\_S} = \pi_{S\_St} [1\text{-}\exp(LR_{0\_S})]$ in the case of model 2.1 (LR_S).

It should be noted that the constraint $0<R_{0\_S}<1$ implies that $0<\delta_{0\_S}<\pi_{S\_St}$ and vice versa, thus leading to a 1:1 correspondence between the two models; however, the upper limits are only theoretical, and have no sense in the context

of non-inferiority studies or clinical trials in general.

Example 4.A.1.1. With π_{S_St} = 0.8, π_{S_Ex} = 0.7, and δ_{0_S} = 0.05, it is possible to obtain the equivalent formulation by calculating R_{0_S} = 1 - 0.05 / 0.8 = 0.9375 in the case of model 2.1 (RS), or by calculating LR_{0_S} = ln(0.9375) = -0.064539 in the case of model 2.2 [ln(R)].

Alternatively, from π_{S_St}=0.8 and R_{0_S}=0.9375 or LR_{0_S}=-0.06454, it is straightforward to calculate δ_{0_S} = 0.05 by inverting the above equation as: $\delta_{0\_S} = \pi_{S\_St}(1 - R_{0\_S})$ or $\delta_{0\_S} = \pi_{S\_St}[1-\exp(LR_{0\_S})]$.

In this way, the formulated hypotheses of model 1 (D_S), model 2.1 (R_S) and model 2.2 (LR_S) are equivalent and, furthermore, as $LR_{0\_S}=\ln(R_{0\_S})$, it is extremely simple to switch from model 2.1 (R_S) to model 2.2 (LR_S), and vice versa.

4.A.1.2 Model 1 (D_S) vs model 3 (odds ratio: OR_S and $\ln(OR_S)$)

Let us consider the following chain of inequalities:

$$\pi_{S\_Ex} - \pi_{S\_St} \le -\delta_0 \rightarrow \pi_{S\_Ex} \le \pi_{S\_St} - \delta_{0\_S} \rightarrow 1-\pi_{S\_Ex} \ge 1-(\pi_{S\_St} - \delta_{0\_S}) \rightarrow$$

$$\rightarrow \frac{\pi_{S\_Ex}}{1-\pi_{S\_Ex}} \le \frac{\pi_{S\_St} - \delta_{0\_S}}{1-(\pi_{S\_St} - \delta_{0\_S})} \rightarrow \frac{\pi_{S\_Ex}}{1-\pi_{S\_Ex}} \cdot \frac{1-\pi_{S\_St}}{\pi_{S\_St}} \le \frac{\pi_{S\_St} - \delta_{0\_S}}{1-(\pi_{S\_St} - \delta_{0\_S})} \cdot \frac{1-\pi_{S\_St}}{\pi_{S\_St}}$$

By putting:

$$OR_{T\_S} = \frac{\pi_{S\_Ex}}{1-\pi_{S\_Ex}} \cdot \frac{1-\pi_{S\_St}}{\pi_{S\_St}}$$

and:

$$OR_{0\_S} = \frac{\pi_{S\_St} - \delta_{0\_S}}{1-(\pi_{S\_St} - \delta_{0\_S})} \cdot \frac{1-\pi_{S\_St}}{\pi_{S\_St}} = \left(1 - \frac{\delta_{0\_S}}{\pi_{S\_St}}\right) \frac{1-\pi_{S\_St}}{1-(\pi_{S\_St} - \delta_{0\_S})}$$

we can obtain the formulation of the statistical hypotheses in terms of the OR. In the case of H_0, we have:

$$H_0 : \pi_{S\_Ex} - \pi_{S\_St} \le -\delta_{0\_S} \quad \rightarrow \quad H_0 : OR \le OR_{0\_S}$$

Otherwise, from model 3, we can consider the hypotheses for model 1 by putting:

$$\pi_{S\_Ex} = \frac{OR_{T\_S}\pi_{S\_St}}{1+\pi_{S\_St}(OR_{T\_S}-1)} \quad \delta_{0\_S} = \frac{\pi_{S\_St}(OR_{0\_S}-1)(1-\pi_{S\_St})}{1+\pi_{S\_St}(OR_{0\_S}-1)}$$

It should be noted that the constraint $0< OR_{0\_S}<1$ implies $0<R_{0\_S}<1$ and vice versa, and so there is a 1:1 correspondence between the two models in these intervals. Furtherly, the same applies in the case of the logarithm transformation:

$$LOR_{0\_S} = \ln(OR_{0\_S}) = \ln(R_{0\_S}) + \ln(1-\pi_{S\_St}) - \ln(1-\pi_{S\_Ex})$$

Once again, it has to be noted that the upper limits are only theoretical, and have no sense in the context of non-inferiority studies or clinical trials in general.

Example 4.A.1.2. With π_{S_St} = 0.8, π_{S_Ex} = 0.7, and δ_{0_S} = 0.05, it is possible to obtain the equivalent formulation for model 3 (OR_S) by calculating:

OR_{0_S} = (1 - 0.05 / 0.8) · (1 - 0.8) / (1 - 0.8 + 0.05) = 0.75 and LOR_{0_S} = ln(0.75)= -0.287682.

Alternatively, from model 3 with π_{S_St} = 0.8, OR_{T_S} = 0.58(3), and OR_{0_S} = 0.75, it is possible to obtain the equivalent formulation for model 1 by calculating:

$\pi_{S\_Ex} = (OR_{T\_S} \cdot \pi_{S\_St}) / (1 + \pi_{S\_St} \cdot (ORT\_S - 1)) = (0.58(3) \cdot 0.8) / (1+0.8 \cdot (0.58(3)-1)) = 0.7$ and, finally:
$\delta_{0\_S} = -\pi_{S\_St} \cdot (OR_{0\_S} - 1) \cdot (1 - \pi_{S\_St}) / (1 + \pi_{S\_St} \cdot (OR_{0\_S} - 1)) = 0.8 \cdot (0.75-1) \cdot (1-0.8) / (1+0.8 \cdot (0.75 - 1)) = 0.05$.

4.A.1.3 Models 2.1 (ratio: R_S) and 2.2 (ln(ratio, LR_S) vs model 3 (odds ratio: OR_S and ln(OR_S))

Let us consider the following chain of inequalities:

$$R_{r\_S} \le R_{0\_S} \rightarrow \pi_{S\_Ex} \le R_{0\_S}\pi_{S\_St} \text{ and } 1-\pi_{S\_Ex} \ge 1-R_{0\_S}\pi_{S\_St} \rightarrow$$
$$\rightarrow \frac{\pi_{S\_Ex}}{1-\pi_{S\_Ex}} \le \frac{R_{0\_S}\pi_{S\_St}}{1-R_{0\_S}\pi_{S\_St}} \rightarrow \frac{\pi_{S\_Ex}}{1-\pi_{S\_Ex}} \cdot \frac{1-\pi_{S\_St}}{\pi_{S\_St}} \le \frac{R_{0\_S}\pi_{S\_St}}{1-R_{0\_S}\pi_{S\_St}} \cdot \frac{1-\pi_{S\_St}}{\pi_{S\_St}} \rightarrow$$
$$\rightarrow \frac{\pi_{S\_Ex}}{1-\pi_{S\_Ex}} \cdot \frac{1-\pi_{S\_St}}{\pi_{S\_St}} \le R_{0\_S} \frac{1-\pi_{S\_St}}{1-R_{0\_S}\pi_{S\_St}}$$
$$\text{with: } \frac{\pi_{S\_Ex}}{1-\pi_{S\_Ex}} \cdot \frac{1-\pi_{S\_St}}{\pi_{S\_St}} = OR_{T\_S} \text{ and } R_{0\_S} \frac{1-\pi_{S\_St}}{1-R_0\pi_{S\_St}} = OR_{0\_S}$$

Consequently, we obtain the following H_0 formulation in terms of OR_S:

$$H_0 : R_{T\_S} \le R_{0\_S} \quad \rightarrow \quad H_0 : OR_{T\_S} \le OR_{0\_S}$$

Note that the constraint $0<R_{0\_S}<1$ implies $0<OR_{0\_S}<1$ and vice versa, and so there is 1:1 correspondence between the formulated hypotheses of the two models, with the same consideration applying to the upper limits.

It is therefore straightforward to obtain the statistical hypotheses for model 3.

Example 4.A.1.3.1. With $\pi_{S\_St} = 0.8$, $R_{T\_S} = 0.875$, and $R_{0\_S} = 0.9375$ (model 2.1), it is possible to obtain the equivalent formulation for model 3 (OR_S) by first calculating:

$OR_{T\_S} = 0.875 \cdot 0.2 / (1 - 0.875 \cdot 0.8) = 0.58(3)$
and then:
$OR_{0\_S} = 0.9375 \cdot 0.2 / (1 - 0.9375 \cdot 0.8) = 0.75$.

Of course, OR_{T_S} can also be obtained using the standard formula after having calculated $\pi_{S\_Ex} = R_{T\_S}\pi_{S\_St} = 0.875 \cdot 0.8 = 0.7$, and then: $OR_{T\_S} = 0.7 \cdot 0.2 / (0.3 \cdot 0.8) = 0.58(3)$.

In the case of model 2.2 [LR_S], it is first necessary to obtain $LR_{0\_S} = \ln(0.9375) = -0.064539$, and then $LOR_{0\_S} = -0.064539 \cdot 0.2 / (1 - 0.9375 \cdot 0.8) = 0.75$.

Furthermore, from model 3 to models 2.1 (R_S) and 2.2 (LR_S), the following formulae apply:

$$R_{T\_S} = \frac{OR_{T\_S}}{\pi_{S\_St}(OR_{T\_S} - 1)} \text{ and } R_{0\_S} = \frac{OR_{0\_S}}{\pi_{S\_St}(OR_{0\_S} - 1)}$$

Example 4.A.1.3.2. With $\pi_{S\_St} = 0.8$, $OR_{T\_S} = 0.58(3)$, and $OR_{0\_S} = 0.75$, it is possible to obtain the equivalent formulation for model 2.1 by calculating:

$R_{T\_S} = OR_{T\_S} / (1 + \pi_{S\_St} \cdot (OR_{T\_S} - 1)) = 0.58(3) / (1 + 0.8 \cdot (0.58(3) - 1)) = 0.857$
and $R_{0\_S} = OR_{0\_S} / (1 + \pi_{S\_St} \cdot (OR_{0\_S} - 1)) = 0.75 / (1 + 0.8 \cdot (0.75 - 1)) = 0.9375$.

So, within these settings, the formulated hypothesis in terms of odds ratios is equivalent to that formulated in terms of relative ratios. Furthermore, the same applies to their log transformations with $LOR_{T\_S} = \ln(OR_{T\_S})$ and $LOR_{0\_S} = \ln(OR_{0\_S})$ or, more explicitly, $LOR_{0\_S} = \ln(R_{0\_S}) + \ln(1 - \pi_{S\_St}) - \ln(1 - R_{0\_S}\pi_{S\_St})$.

4.A.2. Comparison of the sample sizes calculated using the models

Comparisons of the sample sizes calculated using the three models with α and $1-\beta$ fixed require the success probability values, the non-inferiority margin, and the method used to estimate the probability of variance under H_0. Unfortunately, it is

not possible to provide a universally valid rule for choosing the approach leading to the smallest sample size, but we have found a pattern of inequalities in non-inferiority margins that is valid asymptotically and over a clinically relevant interval (see below).

However, it is very easy to calculate sample sizes using a computer program that implements all of the formulae, and therefore choose the most parsimonious approach. To this end, it is useful to use sample size curves in function of π_{S_Ex}, at fixed values of statistical significance (α), power (1-β), π_{S_St}, and non-inferiority margins for the different methods of estimating probabilities.

4.A.2.1. Theoretical results of the sample sizes obtained using all of the models together: 1 (D_S), 2.1 (R_S), 2.2 (LR_S), and 3 (ORS: $LOR_S = \ln(OR_S)$).

A. Comparison of model 1 (D_S) and models 2.1 (R_S) and 2.2 (LR_S)

Models 2.1 and 2.2 require substantially the same number of patients, and are always less demanding (sometimes much less demanding) than model 1, as has also been shown by Laster et al. [49] (see demonstration in Appendix 3.A.2) of the supplementary material.

Example 4.A.2.1. With $\alpha = 0.05$ one-sided, power = 1-$\beta = 0.90$, $\pi_{S\_St} = 0.8$, $\pi_{S\_Ex} = 0.7$, and $\delta_0 = 0.05$, we obtain: $n_{DS\_Ex} = 1,267.45 \approx 1,267$. Alternatively, in the case of model 2.1 with an equivalent non-inferiority margin of $R_0 = (1 - 0.05) / 0.8 = 0.9375$, we obtain: $n_{RS\_Ex} = 1,201.03 \approx 1,201$ using Farrington & Manning's method 1 [38].

In addition, in the case of model 2.2 with $LR_0 = \ln(0.9375) = -0.064539$, we obtain: $n_{LRS\_Ex} = 1,220.83$ 1,221. In the case of method 2, we obtain 1,274.66 (≈1,275) and 1,204.99 (≈1,205), with 1211.59 (≈1212) for nLR_{S_Ex}. Finally, in the case of method 3, we obtain 1,271.40 (≈1,271) and 1,200.43 (≈1,200) with 1,202.50 (≈1,203) for n_{LRS_Ex}.

Furthermore, using method 3 and $\delta_0 = 0.05$, the sample size calculated for the R_S model is about 88% of that calculated for D_S when $\pi_{S\_St} = 0.4$, 90% if $\pi_{S\_St} = 0.5$, and 96% if $\pi_{S\_St} = 0.9$. Finally, using method 3, the sample sizes for the LR_S model are always less than those calculated for the D_S model, and always a little more than those calculated for the R_S model.

B. Asymptotic behaviour study

When π_{S_Ex} under H_A tends to its lower limit, which is $\pi_{S\_St} - \delta_0$ in the case of model 1 (D_S), or when the sample sizes tend to $+\infty$ at a fixed k, non-inferiority margin and πS_St, the following chains of inequalities are valid (see Appendix 3.A.3, particularly the final paragraph Conclusions; see the Appendix 3.B.3 , particularly the final paragraph Conclusions, for failure probabilities) of the supplementary material.

B.1) $n_{RS\_Ex} = n_{LRS\_Ex} < n_{DS\_Ex}$ and $n_{RS\_E}x = n_{LRS\_Ex} < n_{LORS\_Ex}$, regardless of the value of π_{S_St};

B.2) $n_{RS\_Ex} = n_{LRS\_Ex} < n_{LORS\_Ex} \leq n_{DS\_Ex}$ when $\pi_{S\_St} \leq (1+\delta_0)/2$, or $n_{RS\_Ex} = n_{LRS\_Ex} < n_{DS\_Ex} < n_{LORS\_Ex}$ when $\pi_{S\_St} > (1+\delta_0)/2$.

C. Graphical comparisons of the sample sizes obtained using the models

Further results can be obtained using sample size curves for fixed values of the other parameters and varying values of π_{S_Ex} over the clinically relevant interval, with the limits given by the extreme values of non-inferiority ($\pi_{S\_St} - \delta_0$) and π_{S_St}, corresponding to the equipoise condition.

We only show the sample size curves for $\pi_{S\_St} = 0.3$, $\pi_{S\_St} = 0.5$ and $\pi_{S\_St} = 0.7$, with δ_0=0.15 (because a large non-inferiority margin provides a better vision of the sample size curves of the three models) for $\alpha = 0.05$, 1-$\beta = 0.80$, using method 3. In the case of π_{S_St}=0.3, in addition to the fairly parallel pattern of the sample size curves, it is possible to see that: (i) the sample sizes of the D_S model are the largest, with those of the LOR_S model becoming very similar (about 96%) to the values of the D_S model if π_{S_St}= 0.5, and larger if π_{S_St}= 0.7; (ii) the sample sizes of the LOR_S model are always more than those of the R_S and LR_S models; and (iii) the sample sizes of the $R_S$and D_S models become very similar at the highest values of π_{S_St}.

Figures A.1, A.2, and A.3 show that, in addition to being asymptotically valid, the structure of the inequalities is valid over a clinical relevant interval with the equalities being replaced by approximations.

In addition, the pattern of relationships remains substantially the same if the non-inferiority margins are changed; what

FIGURE A.1. $\pi_{S\_St} = 0.3$

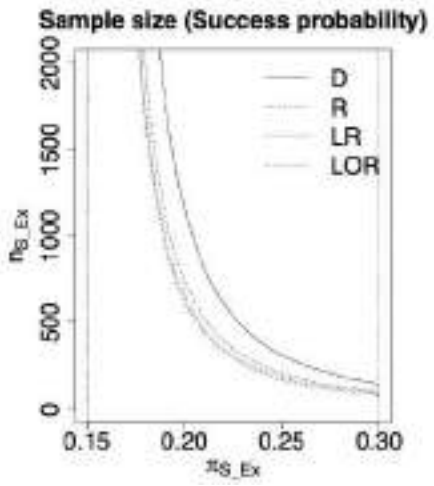


FIGURE A.2. $\pi_{S\_St} = 0.5$

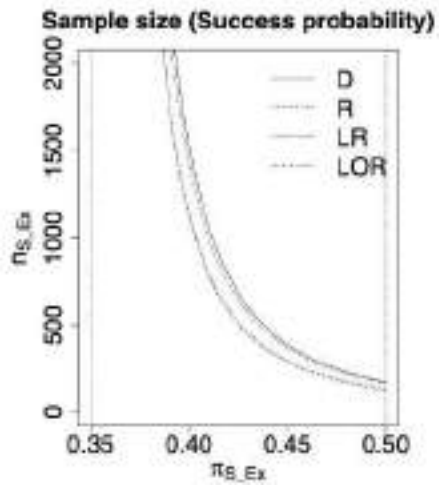
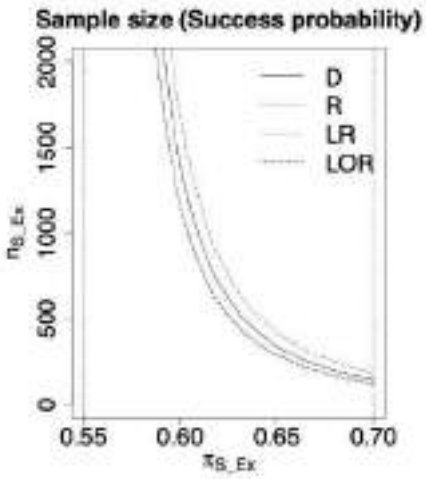


FIGURE A.3. $\pi_{S\_St} = 0.7$



changes is the entity of the differences in sample sizes.

Finally, changing the methods of estimation does not lead to any evident changes in the relationships except in the case of the R_S and LR_S models for which $n_{RLS\_Ex} < n_{RS\_Ex}$ in the case of methods 1 and 2, and $n_{RLS\_Ex} > n_{RS\_Ex}$ in the case of method 3, but the differences are only of a few units.

5. SWITCHING FROM SUCCESS TO FAILURE PROBABILITIES, AND VICE VERSA

In order to be able to enrol as few patients as possible and ensure the most favourable parameterisation, it must be possible to consider that the primary outcome of the trial might be negative, which can be done using the same approach and assumptions as those used when switching from one model to another.

5.1 Hypotheses and non-inferiority margins

Table 3 shows the conversion formulae for switching from success to failure probabilities. For example, starting from $\pi_{S\_St} = 0.8$, $\pi_{S\_Ex} = 0.8$, and $R_{O\_S} = 0.9375$, the non-inferiority margins for the success probability are $\delta_{O\_S} = 0.05$, LR_{O_S}=-0.064538, and LOR_{O_S}=-0.287682. Using the conversion formulae with the complement to 1 of these probabilities ($\pi_{F\_St} = 0.2$, and $\pi_{S\_Ex} = 0.2$) and $\delta_{O\_F} = 0.05$, the non-inferiority margins for the failure probability are R_{O_F}=1.25, LR_{O_F}=0.223143, and LOR_{O_F}=0.287682.

5.2 Sample sizes for the different success and failure models

Comparing the sample size formulae for the models of success and failure, we have:

5.2.1 Model 1: D_S and D_F

It is possible to demonstrate that $n_{DS\_Ex} = n_{DF\_Ex}$ because, at their respective sample sizes, the following expressions at the numerator:

$$\pi_{S\_Ex}(1-\pi_{S\_Ex})=(1-\pi_{F\_Ex})(1-1+\pi_{F\_Ex})=\pi_{F\_Ex}(1-\pi_{F\_Ex})$$
$$\pi_{S\_St}(1-\pi_{S\_St})=(1-\pi_{F\_St})(1-1+\pi_{F\_St})=\pi_{F\_St}(1-\pi_{F\_St})$$

and at the denominator:

$$(\pi_{S\_Ex\_H_A}-\pi_{S\_St\_H_A}+\delta_0)^2=\left[-(\pi_{F\_Ex\_H_A}-\pi_{F\_St\_H_A}-\delta_0)\right]^2=$$
$$=(\pi_{F\_Ex\_H_A}-\pi_{F\_St\_H_A}-\delta_0)^2$$

are the same.

5.2.2 Model 2: R_S, LR_S, and R_F, LR_F

In the case of the models R_S and R_F the denominators of the sample size formulae are equal, but the numerators are different because $R_{O\_S} \leq 1 \leq R_{O\_F}$. Therefore, except in the non-sensible case of $R_{O\_F} = R_{O\_S} = 1$, we have $n_{RS\_Ex} < n_{RF\_Ex}$, which means that a success-based approach is preferable.

In the case of the LR models, given that $n_{LRS\_Ex} \sim n_{RS\_Ex}$ and $n_{LRF\_Ex} \sim n_{RF\_Ex}$, it can be concluded that $n_{LRS\_Ex} < n_{LRF\_Ex}$

5.2.3 Model 3 (LOR_S) and OR_F (LOR_F)

It is possible to demonstrate that $n_{LORF\_Ex} = n_{LORS\_Ex}$ because, by definition:

$OR_{S\_T} = 1/OR_{F\_T}$ and $OR_{S\_0} = 1/OR_{F\_0}$

and, consequently, $n_{LORF\_Ex} = n_{LORS\_Ex}$.

5.3 CONCLUSIONS

In the case of models (D) and OR (LOR), the sample sizes are equal regardless of whether we are considering the success or failure probability; in the case of model R (and LR), lower sample sizes are obtained by using success probabilities.

Given the sample sizes calculated in function of δ_{S_Ex} varying over the clinically relevant interval, the following inequality chains apply:

i) if $\pi_{S\_st} \leq (1+\delta_0)/2$, $n_{RS\_Ex} <^* n_{LRS\_Ex} < n_{DS\_Ex} = n_{DF\_Ex} <^* n_{LORS\_Ex} = n_{LORF\_Ex} <^* n_{RF\_Ex} <^* n_{LRF\_Ex}$
ii) if $\pi_{S\_st} > (1+\delta_0)/2$, $n_{RS\_Ex} <^* n_{LRS\_Ex} < n_{LORS\_Ex} = n_{LORF\_Ex} < n_{DS\_Ex} = n_{DF\_Ex} < n_{RF\_Ex} <^* n_{LRF\_Ex}$.

(<* means that the difference is only a few units).

It has to be stressed that the above inequalities (with model R_S as the best followed by LR_S) come from method 3 (constrained MLE), which performs better than method 1 of Blackwelder [35] or method 2 of Dunnett and Gent [3] in terms of controlling the type I error probability, power, and confidence interval coverage. Furthermore, the difference between these two models is practically eliminated using method 2, and reversed using method 1 (model LRS is the best, even if by only a few units).

Example 5.3.1. A case with $\pi_{S\_St} > (1+\delta_0)/2$

Using method 3, if α =0.05, β =0.2 and k=1, π_{S_St} = 0.7, π_{S_Ex} = 0.6, and δ_{O_S} = 0.15 for the success probability, and π_{F_St} = 0.3, π_{F_Ex} = 0.4 and δ_{O_F} = 0.15, for the failure probability.

The table below shows the sample sizes for the four models.

	n_D	n_R	n_{LR}	n_{LOR}
Success	1,105.047~1,105	914.107~914	924.168~924	1,331.724~1,332
Failure	1,105.047~1,105	1,733.555~1,734	1,753.843~1,754	1,331.724~1,332

Figures 5.3.1.1 and 5.3.1.2 show the sample size curves relating to the example above, which confirm the inequality chains for the success and failure probabilities separately.

FIGURE 5.3.1.1. Sample size curves of the Example 5.3.1.

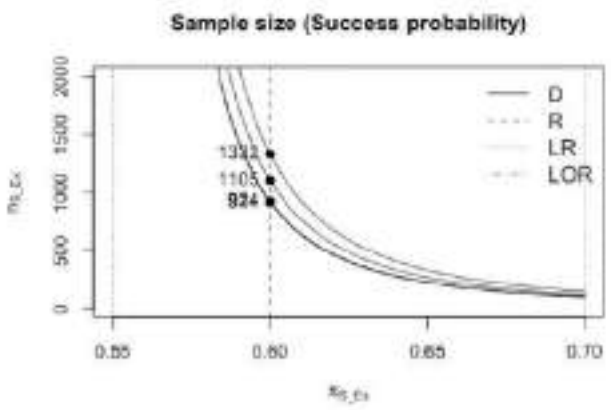
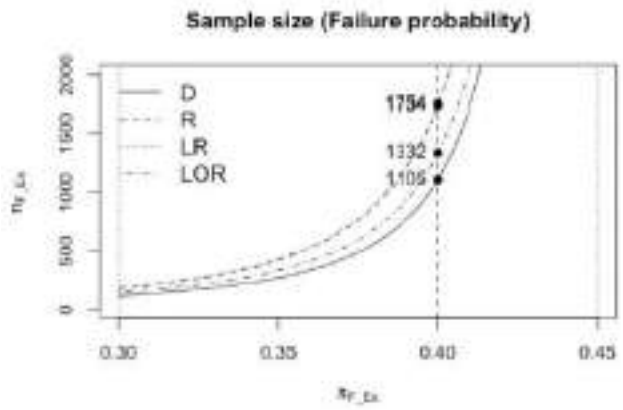


FIGURE 5.3.1.2. Sample size curves of the Example 5.3.1.



Example 5.3.2. A case with $\pi_{S\_St}<(1+\delta_0)/2$

Using method 3, if $\alpha = 0.05$, $\beta = 0.2$ and $k = 1$, $\pi_{S\_St} = 0.4$, $\pi_{S\_Ex} = 0.3$, and $\delta_{0\_S} = 0.15$ for the success probability, and $\pi_{F\_St} = 0.6$, $\pi_{F\_Ex} = 0.7$ and $\delta_{0\_F} = 0.15$, for the failure probability.

The table below shows the sample sizes for the four models:

	n_D	n_R	n_{LR}	n_{LOR}
Success	1,105.047~1,105	730.199~730	745.526~746	887.249~887
Failure	1,105.047~1,105	1,446.498~1,446	1,457.99~1,458	887.249~887

Figures 5.3.2.1 and 5.3.2.2 show the sample size curves relating to the example above, which confirm the inequality chains for the success and failure probabilities separately.

FIGURE 5.3.2.1. Sample size curves of the Example 5.3.2.

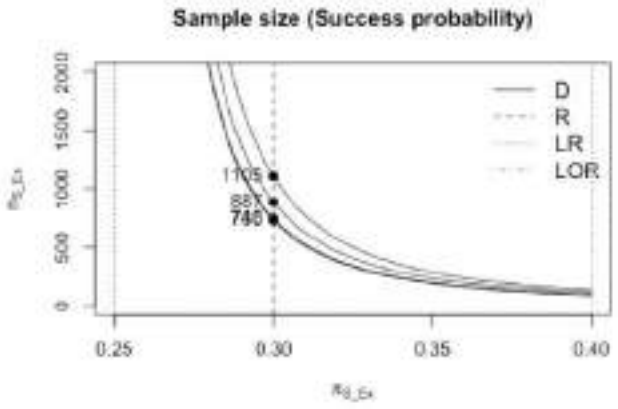
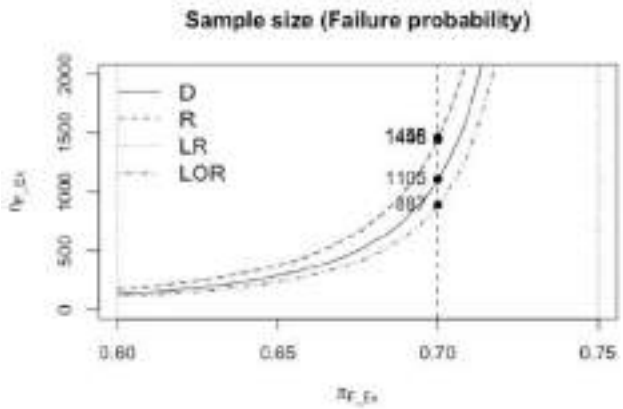


FIGURE 5.3.2.2. Sample size curves of the Example 5.3.2.



DISCUSSION

Biomedical research has to be adequately empowered by appropriate sample sizes for economic, organisational,

logistic, scientific and, mainly, ethical reasons (even if it is practically impossible to separate the ethical and scientific aspects of biomedical research). In addition, the feasibility of a trial mainly depends on the sample size that it is possible to enrol. It is important that a research study is adequately powered with the smallest possible number of subjects, particularly in non-inferiority settings in which it is easy to increase the non-inferiority margin in order to obtain a smaller sample size. However, a too small sample size in a non-inferiority setting not only fails to demonstrate that the experimental drug is non-inferior to the optimal standard treatment, but also fails to demonstrate that it is superior to a placebo or a previous standard as it can be taken granted that the demonstrated non-inferiority of the experimental drug also leads to the demonstrated superiority of the experimental drug over a placebo or the previous standard.

The search for a statistical approach that leads to the most parsimonious but adequate sample size is particularly important when comparing two probabilities for which different parameterisation models, testing procedure and sample size calculation formulae are available. We have shown the statistical models, the methods of estimating variance under H_0, and sample size calculation formulae separately for success and failure probabilities, and described a method of consistently switching among the models and probabilities in order to choose the most parsimonious approach. To this purpose, the coherency of the formulations is kept by the general constraint that the NI margin of the final model is calculated by using only the true probability of the standard (π_{St}) which is independent of the non-inferiority NI margin and considered "known" in the phase of planning the study.

We have also demonstrated that, asymptotically, there is a hierarchical structure of inequalities among the sample sizes of the different models, and verified that it does not change under H_A within the range of clinically plausible values for non-inferiority settings.

We confirm that the sample sizes for the R_S model are smaller in the case of success probabilities as has been previously shown by Laster et al. [49]. However, it has to be pointed out that the greater efficiency of the R_S model is not maintained in the case of failure probabilities, for which the sample size of the R_F model is greater than that of the D_F model, a result that it is the opposite of that described by Laster et al. [49]. It has to be said that Laster et al. [49] obtained their result by reversing the order of the ratio between Experimental and Standard used in the case of success probability in order to ensure that it remained less than 1. However, this reversal and a different formulation of H_0 does not lead to a single inferiority margin because it depends on π_{F_Ex}, as shown in Appendix 1.2.

We have also shown that each success model has an equivalent model for failure. In the case of the D_S and D_F and LOR_S and LOR_F models, the sample sizes are the same, whereas. the sample sizes of the R_S and LR_S models are always smaller than those of the R_F and LR_F models. It is thus possible to establish a hierarchical structure of sample sizes for the eight equivalent models within the clinically relevant interval of π_{S_Ex} (π_{F_Ex}) under H_A when all of the other parameters are fixed.

Furthermore, the odds ratio model leads to a larger sample size and, consequently, is not to be preferred even if an effect size or a non-inferiority margin expressed on the basis of this parameterisation might seem to be sensible.

The most sensible approach is to consider each case separately by calculating the pertinent sample size curves over the pertinent interval of clinical non-inferiority by using the usual parametrization in a particular clinical setting, and then choosing the one that leads to the most parsimonious sample size.

TABLE 1. Success Probability. Null Hypothesis (H_0) of the three considered Models (M), together with their sampling distribution, and sample size calculation formulae for the experimental group.

	H_{0_S}	Sampling Distribution	$n(\ )_{S\_Ex} = \cdots \quad n(\ )_{S\_St} = k \cdot n(\ )_{S\_Ex}$
M. 1	$\pi_{S\_Ex} - \pi_{S\_St} \le -\delta_{S\_S}$ with $\delta_{S\_S} > 0$	$D_S = p_{S\_Ex} - p_{S\_St} + \delta_{S\_S}$; $\mu_{D_S} = \pi_{S\_Ex} - \pi_{S\_St} + \delta_{S\_S}$; $\sigma^2_{D_S} = \frac{\pi_{S\_Ex}(1-\pi_{S\_Ex})}{n_{D_{S\_Ex}}} + \frac{\pi_{S\_St}(1-\pi_{S\_St})}{n_{D_{S\_St}}}$	$n_{D_S} = \frac{\left[z_{1-\alpha}\sqrt{\pi_{S\_Ex\_H_0}(1-\pi_{S\_Ex\_H_0}) + \frac{\pi_{S\_St\_H_0}(1-\pi_{S\_St\_H_0})}{k}} + z_{1-\beta}\sqrt{\pi_{S\_Ex\_H_A}(1-\pi_{S\_Ex\_H_A}) + \frac{\pi_{S\_St\_H_A}(1-\pi_{S\_St\_H_A})}{k}}\right]^2}{(\pi_{S\_Ex\_H_A} - \pi_{S\_St\_H_A} + \delta_{S\_S})^2}$
M. 2.1	$R_{S\_S} \le R_{S\_S}$; $R_{S\_S} = \frac{\pi_{S\_Ex}}{\pi_{S\_St}}$ and $0 < R_{S\_S} \le 1$ or $\pi_{S\_Ex} - R_{S\_S}\pi_{S\_St} \le 0$	$R_S = p_{S\_Ex} - R_{S\_S}\, p_{S\_St}$; $\mu_{R_S} = \pi_{S\_Ex} - R_{S\_S}\,\pi_{S\_St}$; $\sigma^2_{R_S} = \frac{\pi_{S\_Ex}(1-\pi_{S\_Ex})}{n_{R_{S\_Ex}}} + R^2_{S\_S}\frac{\pi_{S\_St}(1-\pi_{S\_St})}{n_{R_{S\_St}}}$	$n_{R_{S\_S}} = \frac{\left[z_{1-\alpha}\sqrt{\pi_{S\_Ex\_H_0}(1-\pi_{S\_Ex\_H_0}) + R^2_S\frac{\pi_{S\_St\_H_0}(1-\pi_{S\_St\_H_0})}{k}} + z_{1-\beta}\sqrt{\pi_{S\_Ex\_H_A}(1-\pi_{S\_Ex\_H_A}) + R^2_S\frac{\pi_{S\_St\_H_A}(1-\pi_{S\_St\_H_A})}{k}}\right]^2}{(\pi_{S\_Ex\_H_A} - R_S\,\pi_{S\_St\_H_A})^2}$
M. 2.2	$LR_{S\_S} \le LR_{S\_S}$ with $R_{Y\_S} = \ln(R_{Y\_S})$ and $LR_{S\_S} \le 0$	$LR_S = \ln\left(\frac{p_{S\_Ex}}{p_{S\_St}}\right) - LR_{S\_S}$; $\mu_{LR_S} = \ln\left(\frac{\pi_{S\_Ex}}{\pi_{S\_St}}\right) - LR_{S\_S} = \ln(R_{S\_S}) - \ln(R_{S\_S})$; $\sigma^2_{LR_S} = \frac{1-\pi_{S\_Ex}}{n_{LR_{S\_Ex}} \cdot \pi_{S\_Ex}} + \frac{1-\pi_{S\_St}}{n_{LR_{S\_St}} \cdot \pi_{S\_St}}$	$n_{LR_{S\_S}} = \frac{\left[z_{1-\alpha/2}\sqrt{\frac{1-\pi_{S\_Ex\_H_0}}{\pi_{S\_Ex\_H_0}} + \frac{1-\pi_{S\_St\_H_0}}{k \cdot \pi_{S\_St\_H_0}}} + z_{1-\beta}\sqrt{\frac{1-\pi_{S\_Ex\_H_A}}{\pi_{S\_Ex\_H_A}} + \frac{1-\pi_{S\_St\_H_A}}{k \cdot \pi_{S\_St\_H_A}}}\right]^2}{\left(\ln\left(\pi_{S\_Ex\_H_A}/\pi_{S\_St\_H_A}\right) - LR_{S\_S}\right)^2}$
M. 3	$OR_{S\_S} \le OR_{S\_S}$ with $OR_{S\_S} = \frac{\pi_{S\_Ex}}{1-\pi_{S\_Ex}} / \frac{\pi_{S\_St}}{1-\pi_{S\_St}}$ and $0 < OR_{S\_S} \le 1$ or $LOR_{S\_S} \le LOR_{S\_S}$ with $LOR_{S\_S} = \ln(OR_{S\_S})$ and $LOR_{S\_S} = \ln(OR_{S\_S}) \le 0$	$LOR_S = \ln(OR_S) - \ln(OR_{S\_S}) = \ln(p_{S\_Ex}) - \ln(1-p_{S\_Ex}) - \ln(p_{S\_St}) + \ln(1-p_{S\_St}) - LOR_{S\_S}$; $\mu_{LOR_S} = \ln(OR_{S\_S}) - LOR_{S\_S}$; $\sigma^2_{LOR_S} = \frac{1}{n_{LOR_{S\_Ex}} \cdot \pi_{S\_Ex}(1-\pi_{S\_Ex})} + \frac{1}{n_{LOR_{S\_St}} \cdot \pi_{S\_St}(1-\pi_{S\_St})}$	$n_{LOR_{S\_S}} = \frac{\left[z_{1-\alpha/2}\sqrt{\frac{1}{\pi_{S\_Ex\_H_0}(1-\pi_{S\_Ex\_H_0})} + \frac{1}{k \cdot \pi_{S\_St\_H_0}(1-\pi_{S\_St\_H_0})}} + z_{1-\beta}\sqrt{\frac{1}{\pi_{S\_Ex\_H_A}(1-\pi_{S\_Ex\_H_A})} + \frac{1}{k \cdot \pi_{S\_St\_H_A}(1-\pi_{S\_St\_H_A})}}\right]^2}{\left(\ln(\pi_{S\_Ex\_H_A}) - \ln(1-\pi_{S\_Ex\_H_A}) - \ln(\pi_{S\_St\_H_A}) + \ln(1-\pi_{S\_St\_H_A}) - LOR_{S\_S}\right)^2}$

Table 1.1 Success Probability. Sample sizes for α = 0.025 and 1-β = 0.80

π_{S_St}	R_{T_S}	M	0.85 D_S	0.85 R_S	0.85 LRS	0.85 LOR_S	0.9 D_S	0.9 R_S	0.9 LRS	0.9 LOR_S	0.95 D_S	0.95 R_S	0.95 LRS	0.95 LOR_S	1.0 D_S	1.0 R_S	1.0 LRS	1.0 LOR_S
0.1	0.8	1	52,674	42,502	42,200	43,091	13,932	10,946	10,813	11,040	6,139	5,008	4,924	5,035	3,532	2,896	2,837	2,903
0.1		2	52,655	42,305	42,198	43,249	13,884	10,851	10,883	11,118	6,136	4,945	4,969	5,072	3,528	2,849	2,865	2,925
0.1		3	52,741	42,296	42,410	43,249	13,926	10,848	10,899	11,116	6,161	4,942	4,973	5,072	3,549	2,847	2,868	2,923
0.1	0.85	1	.	.	.	.	53,969	46,128	45,913	46,053	33,812	11,852	11,754	11,951	6,279	5,408	5,369	5,441
0.1		2	.	.	.	.	53,955	45,981	46,043	46,766	33,807	11,780	11,808	11,995	6,276	5,361	5,376	5,463
0.1		3	.	.	.	.	54,018	45,974	46,057	46,766	33,837	11,777	11,813	11,995	6,296	5,359	5,380	5,463
0.1	0.95	1	-	-	-	-	-	-	-	-	.	.	.	.	56,511	53,757	53,888	53,999
0.1		2	-	-	-	-	-	-	-	-	.	.	.	.	56,509	53,710	53,726	54,023
0.1		3	-	-	-	-	-	-	-	-	.	.	.	.	56,529	53,708	53,729	54,023
0.3	0.8	1	13,953	11,315	11,222	12,135	3,550	2,891	2,850	3,090	1,604	1,311	1,287	1,399	936	751	736	802
0.3		2	13,933	11,265	11,205	12,162	3,542	2,867	2,877	3,105	1,599	1,298	1,302	1,408	934	741	745	806
0.3		3	13,944	11,259	11,299	12,162	3,547	2,864	2,884	3,105	1,602	1,294	1,307	1,408	935	739	749	806
0.3	0.85	1	.	.	.	.	14,201	12,168	12,101	12,853	3,609	3,100	3,072	3,271	1628	1,402	1,387	1475
0.3		2	.	.	.	.	14,188	12,131	12,146	12,879	3,603	3,084	3,090	3,283	1625	1,393	1,396	1484
0.3		3	.	.	.	.	14,195	12,127	12,157	12,879	3,606	3,081	3,096	3,283	1627	1,390	1,400	1484
0.3	0.95	1	-	-	-	-	-	-	-	-	.	.	.	.	1,4651	13,957	13,922	14,223
0.3		2	-	-	-	-	-	-	-	-	.	.	.	.	1,4648	13,927	13,933	14,228
0.3		3	-	-	-	-	-	-	-	-	.	.	.	.	1,4651	13,925	13,935	14,228
0.5	0.8	1	6,208	5,078	5,025	5,565	1,562	1,279	1,257	1,505	697	571	560	674	392	322	315	382
0.5		2	6,189	5,057	5,063	5,565	1,554	1,270	1,275	1,513	692	566	569	679	390	319	321	385
0.5		3	6,199	5,056	5,081	5,565	1,554	1,270	1,283	1,513	692	566	575	679	390	319	326	385
0.5	0.85	1	.	.	.	.	6,148	5,328	5,199	6,112	1,568	1,310	1,286	1,538	688	601	594	687
0.5		2	.	.	.	.	6,134	5,362	5,306	6,126	1,562	1,344	1,347	1,543	695	588	600	690
0.5		3	.	.	.	.	6,134	5,361	5,381	6,126	1,562	1,344	1,354	1,543	695	588	605	690
0.5	0.95	1	-	-	-	-	-	-	-	-	.	.	.	.	6,279	5,573	5,966	6,269
0.5		2	-	-	-	-	-	-	-	-	.	.	.	.	6,278	5,570	5,972	6,271
0.5		3	-	-	-	-	-	-	-	-	.	.	.	.	6,278	5,570	5,977	6,271
0.7	0.8	1	2,890	2,405	2,369	3,306	730	589	575	839	308	254	248	367	168	138	135	203
0.7		2	2,870	2,397	2,396	3,405	702	586	586	852	304	254	254	375	165	138	139	208
0.7		3	2,874	2,404	2,424	3,405	704	590	601	852	306	257	265	375	167	141	147	208
0.7	0.85	1	.	.	.	.	2,835	2,466	2,441	3,294	693	600	592	801	293	258	255	345
0.7		2	.	.	.	.	2,825	2,460	2,460	3,286	688	599	599	811	290	258	259	354
0.7		3	.	.	.	.	2,829	2,468	2,483	3,286	690	603	611	811	288	261	267	354
0.7	0.95	1	.	.	.	.	.	.	.	.	.	.	.	.	2,691	2,560	2,557	2,861
0.7		2	.	.	.	.	.	.	.	.	.	.	.	.	2,688	2,560	2,561	2,864
0.7		3	.	.	.	.	.	.	.	.	.	.	.	.	2,690	2,564	2,570	2,864
0.9	0.8	1	1,046	820	893	2,353	216	205	196	541	92	78	75	218	66	26	35	111
0.9		2	1,036	919	914	2,486	228	206	205	605	88	80	80	258	41	38	38	136
0.9		3	1,060	951	947	2,486	247	224	233	605	102	93	100	258	53	48	54	136
0.9	0.85	1	.	.	.	.	945	849	830	1,900	207	183	178	426	78	67	66	169
0.9		2	.	.	.	.	932	849	846	2,002	202	186	184	472	75	69	69	193
0.9		3	.	.	.	.	964	881	894	2,002	220	203	210	472	89	82	88	193
0.9	0.95	1	-	-	-	-	-	-	-	-	.	.	.	.	688	664	663	975
0.9		2	-	-	-	-	-	-	-	-	.	.	.	.	695	666	666	978
0.9		3	-	-	-	-	-	-	-	-	.	.	.	.	713	686	682	988

Legend: π_{S_St} = true success probability for the Standard drug, R_{0_S} = non-inferiority margin expressed in the ratio scale, M = Method 1, 2, and 3 (see text); R_{T_S} = true ratio between the true success probability for the Experimental drug (π_{S_Ex}) and π_{S_St}; D_S = Difference, R_S = Ratio, LRS = ln(RS), LOR_S = ln(Odds Ratio). The "–" sign means that it is a case incompatible with non-inferiority and the "." sign means that the denominator of the sample size formula is equal to 0 (sample size tends to infinity)

Table 2. Success Probability. Formulae for switching from a model to another of the three considered models

From the Model:	To the Model:		
Model 1	Model 2.1	Model 2.2	Model 3
$\{\pi_{S\_Ex}, \pi_{S\_St} \rightarrow \delta_{0\_S}$	$R_{T\_S} = \frac{\pi_{S\_Ex}}{\pi_{S\_St}}$ $R_{0\_S} = 1 - \frac{\delta_{0\_S}}{\pi_{S\_St}} = \frac{\pi_{S\_St} - \delta_{0\_S}}{\pi_{S\_St}}$	$LR_{T\_S} = \ln\left(\frac{\pi_{S\_Ex}}{\pi_{S\_St}}\right)$ $LR_{0\_S} = \ln\left(1 - \frac{\delta_{0\_S}}{\pi_{S\_St}}\right)$	$OR_{T\_S} = \frac{\pi_{S\_Ex}}{\pi_{S\_St}} \frac{1-\pi_{S\_St}}{1-\pi_{S\_Ex}}$; $LOR_{T\_S} = \ln(OR_{T\_S})$ $OR_{0\_S} = \left(1 - \frac{\delta_{0\_S}}{\pi_{S\_St}}\right)\left(\frac{1-\pi_{S\_St}}{1-\pi_{S\_St}+\delta_{0\_S}}\right)$; $LOR_{0\_S} = \ln(OR_{0\_S})$
Model 2.1	Model 1	Model 2.2	Model 3
$\{\pi_{S\_St}, R_{T\_S} \rightarrow R_{0\_S}$	$\pi_{S\_Ex} = R_{T\_S} \cdot \pi_{S\_St}$ $\delta_{0\_S} = (1 - R_{0\_S})\pi_{S\_St}$	$LR_{T\_S} = \ln(R_{T\_S})$ $LR_{0\_S} = \ln(R_{0\_S})$	$OR_{T\_S} = R_{T\_S} \frac{1-\pi_{S\_St}}{1-R_{T\_S} \cdot \pi_{S\_St}}$; $LOR_{T\_S} = \ln(OR_{T\_S})$ $OR_{0\_S} = R_{0\_S} \frac{1-\pi_{S\_St}}{1-R_{0\_S} \cdot \pi_{S\_St}}$; $LOR_{0\_S} = \ln(OR_{0\_S})$
Model 2.2	Model 1	Model 2.1	Model 3
$\{\pi_{S\_St}, LR_{T\_S} \rightarrow LR_{0\_S}$	$\pi_{S\_Ex} = \exp(LR_{T\_S}) \cdot \pi_{S\_St}$ $\delta_{0\_S} = [1 - \exp(LR_{0\_S})] \cdot \pi_{S\_St}$	$R_{T\_S} = \exp(LR_{T\_S})$ $R_{0\_S} = \exp(LR_{0\_S})$	$OR_{T\_S} = \exp(LR_{T\_S}) \frac{1-\pi_{S\_St}}{1-\exp(LR_{T\_S}) \cdot \pi_{S\_St}}$; $LOR_{T\_S} = \ln(OR_{T\_S})$ $OR_{0\_S} = \exp(LR_{0\_S}) \frac{1-\pi_{S\_St}}{1-\exp(LR_{0\_S}) \cdot \pi_{S\_St}}$; $LOR_{0\_S} = \ln(OR_{0\_S})$
Model 3	Model 1	Model 2.1	Model 2.2
$\{\pi_{S\_St}, OR_{T\_S} \rightarrow OR_{0\_S}$	$\pi_{S\_Ex} = \frac{OR_{T\_S} \cdot \pi_{S\_St}}{1 + \pi_{S\_St}(OR_{T\_S} - 1)}$ $\delta_{0\_S} = \frac{\pi_{S\_St}(1 - \pi_{S\_St})(OR_{0\_S} - 1)}{1 + \pi_{S\_St}(OR_{0\_S} - 1)}$	$R_{T\_S} = \frac{OR_{T\_S}}{1 + \pi_{S\_St}(OR_{T\_S} - 1)}$ $R_{0\_S} = \frac{OR_{0\_S}}{1 + \pi_{S\_St}(OR_{0\_S} - 1)}$	$LR_{T\_S} = \ln(OR_{T\_S}) - \ln(1 + \pi_{S\_St}(OR_{T\_S} - 1))$ $LR_{0\_S} = \ln(OR_{0\_S}) - \ln(1 + \pi_{S\_St}(OR_{0\_S} - 1))$

Table 3. Formulae for converting the H_0 and H_A success hypotheses into those of failure.

Model	From Success	To Failure	Non-inferiority Margin
1. Difference	$H_0: \pi_{S,0} - \pi_{S,0} \leq -\delta_{0,S}$ $H_A: \pi_{S,0} - \pi_{S,0} > -\delta_{0,S}$ with $\delta_{0,S} > 0$	$H_0: \pi_{F,0} - \pi_{F,0} \geq \delta_{0,F}$ $H_A: \pi_{F,0} - \pi_{F,0} < \delta_{0,F}$ with $\delta_{0,F} > 0$	$\delta_{0,F} = \delta_{0,S}$
2.1 Ratio	$H_0: R_{T,S} \leq R_{0,S}$ $H_A: R_{T,S} > R_{0,S}$ with $R_{S,T} = \frac{\pi_{S,0}}{\pi_{S,0}}$ and $R_{0,S} < 1$	$H_0: R_{T,F} \geq R_{0,F}$ $H_A: R_{T,F} < R_{0,F}$ with $R_{T,F} = \frac{\pi_{F,0}}{\pi_{F,0}}$ and $R_{0,F} > 1$	$R_{0,F} = \frac{1-R_{0,S}(1-\pi_{F,0})}{\pi_{F,0}} = \frac{1-R_{0,S}\pi_{S,0}}{1-\pi_{S,0}}$ or $R_{0,S} = \frac{1-R_{0,F}(1-\pi_{S,0})}{\pi_{S,0}} = \frac{1-R_{0,F}\pi_{F,0}}{1-\pi_{F,0}}$
2.2 Ln(Ratio)	$H_0: LR_{T,S} \leq LR_{0,S}$ $H_A: LR_{T,S} > LR_{0,S}$ with $LR_{T,S} = ln\left(\frac{\pi_{S,0}}{\pi_{S,0}}\right)$ and $LR_{0,S} < 0$	$H_0: LR_{T,F} \geq LR_{0,F}$ $H_A: LR_{T,F} < LR_{0,F}$ with $LR_{T,F} = ln\left(\frac{\pi_{F,0}}{\pi_{F,0}}\right)$ and $LR_{0,F} > 0$	$LR_{0,F} = ln\left(\frac{1-R_{0,S}(1-\pi_{S,0})}{\pi_{F,0}}\right) = ln\left(\frac{1-R_{0,S}\pi_{S,0}}{1-\pi_{F,0}}\right)$ or $R_{0,S} = ln\left(\frac{1-R_{0,F}(1-\pi_{S,0})}{\pi_{S,0}}\right) = ln\left(\frac{1-R_{0,F}\pi_{F,0}}{1-\pi_{F,0}}\right)$
3. Odds Ratio and ln(Odds Ratio)	$H_0: OR_{T,S} \leq OR_{0,S}$ $H_A: OR_{T,S} > OR_{0,S}$ with $OR_{T,S} = \frac{\pi_{S,0}}{1-\pi_{S,0}} \Big/ \frac{\pi_{S,0}}{1-\pi_{S,0}}$ and $OR_{0,S} < 1$ $H_0: LOR_{T,S} \leq LOR_{0,S}$ $H_A: LOR_{T,S} > LOR_{0,S}$ with $LOR_{T,S} = ln(OR_{T,S})$ and $LOR_{0,S} < 0$	$H_0: OR_{T,F} \geq OR_{0,F}$ $H_A: OR_{T,F} < OR_{0,F}$ with $OR_{T,F} = \frac{\pi_{F,0}}{1-\pi_{F,0}} \Big/ \frac{\pi_{F,0}}{1-\pi_{F,0}}$ and $OR_{0,F} > 1$ $H_0: LOR_{T,F} \geq LOR_{0,F}$ $H_A: LOR_{T,F} < LOR_{0,F}$ with $LOR_{T,F} = ln(OR_{T,F})$ and $LOR_{0,F} > 0$	$OR_{0,F} = \frac{1}{OR_{0,S}}$ or: $OR_{0,S} = \frac{1}{OR_{0,F}}$ $LOR_{0,F} = -ln(OR_{0,S})$ or: $LOR_{0,S} = -ln(OR_{0,F})$

References

1. Garattini S, Bertelè V, Li Bassi L. How can research ethics committees protect patients better?. British Medical Journal, 2003; 236:1199-1201.
2. Cesana BM, Marubini E. Re: Further topics for research ethics committees. British Medical Journal http://www.bmj.com/content/326/7400/1199/rapid-responses (accessed on December, 2017)
3. Dunnett CW, Gent M. Significance Testing to Establish Equivalence between Treatments, with Special Reference to Data in the Form of 2 x 2 Tables. Biometrics, 1977; 33(4):593-596.
4. Gart JJ. The comparison of proportions: a review of significance tests, confidence intervals and adjustments for stratification. Review of the International Statistical Institute, 1971;39: 148-169.
5. The European Agency for the Evaluation of Medicinal Products (CPMP) Points to Consider on Biostatistical/Methodological Issues Arising from Recent CPMP Discussions on Licensing Applications: Choice of Delta CPMP/EWP/2158/99 London, 23 September 1999.
6. Committee for Proprietary Medicinal Products (CPMP). Points to Consider on the Choice of Non-Inferiority Margin (Draft). CPMP/EWP/2158/99 London 2004.
7. Committee for Medicinal Products for Human Use (CHMP). Guideline On The Choice Of The Non-Inferiority Margin Doc. Ref. EMEA/CPMP/EWP/2158/99 London, 27 July 2005.
8. Food and Drug Administration (CDER) (CBER) Non-Inferiority Clinical Trials to Establish Effectiveness Guidance for Industry November 2016 Clinical/Medical FDA's guidance "Non-Inferiority Clinical Trials".
9. Developing Antiretroviral Drugs for Treatment Guidance for Industry U.S.Department of Health and Human Services Food and Drug Administration Center for Drug Evaluation and Research (CDER) November 2015 Clinical/Antimicrobial Revision 1 Appendix B: Noninferiority Margin Recommendations and Justifications].
10. Food and Drug Administration Center for Drug Evaluation and Research (CDER) Guidance for Industry Diabetes Mellitus: Developing Drugs and Therapeutic Biologics for Treatment and Prevention Draft Guidance February 2008 Clinical/Medical.

11. Committee for Medicinal Products for Human Use (CHMP) Guideline on clinical investigation of medicinal products in the treatment of diabetes mellitus CPMP/EWP/1080/00 rev. 1. The European Agency for the Evaluation of Medicinal Products (EMEA) Draft 16 September 2011.
12. Committee for Proprietary Medicinal Products (CPMP). Notes for Guidance on the Evaluation of Medicinal Products Indicated for the Treatment of Bacterial Infections. CPMP/EWP/558/95 The European Agency for the Evaluation of Medicinal Products (EMEA) 1997, rev 1. 2004.
13. Food and Drug Administration (FDA). (1992). Points to Consider. Clinical Evaluation of Anti-infective Drug Products.
14. Food and Drug Administration Center for Drug Evaluation and Research (CDER) August 2016 Clinical/Antimicrobial Microbiology Data for Systemic Antibacterial Drugs — Development, Analysis, and Presentation Guidance for Industry.
15. Röhmel J. Statistical considerations of FDA and CPMP rules for the investigation of new anti-bacterial products Statistics in Medicine 2001;20:2561–2571. DOI: 10.1002/sim.729.
16. Holmgren EB. Establishing equivalence by showing that a specified percentage of the effect of the active control over placebo is maintained. J. Biopharm. Stat., 2000; 9: 651–659.
17. D'Agostino RB. Sr., Massaro JM, Sullivan LM. Non-inferiority trials: design concepts and issues – the encounters of academic consultants in statistics. Statist. Med., 2003;22: 169–186. DOI: 10.1002/sim.1425.
18. Snapinn SM. Noninferiority trials. Curr. Control Trials Cardiovasc. Med.M 2000; 1(1): 19–21.
19. Pigeot I, Schäfer J, Röhmel J, Hauschke D. Assessing non-inferiority of a new treatment in a three-arm Clinical trial including a placebo. Statist. Med. 2003; 22: 883-899. DOI:10.1002/sim.l450.
20. Hung HMJ, Wang SJ, O'Neill R. A Regulatory Perspective on Choice of Margin and Statistical Inference Issue in Non-inferiority Trials. Biometrical Journal, 2005; 47(1): 28–36. DOI: 10.1002/bimj.200410084
21. Hung HMJ, Wang SJ, Tsong Y, Lawrence J, O'Neil RT. Some fundamental issues with non-inferiority testing in active controlled trials. Statist. Med. 2003;22:213–225 (DOI: 10.1002/sim.1315)]
22. Wiens BL. Choosing an equivalence limit for noninferiority or equivalence studies, Controlled Clinical Trials 23 (2002) 2–14. Erratum in Controlled Clinical Trials 2002 Dec; 23(6):774
23. Tsong Y, Higgins K, Wang SJ, Hung HMJ. An overview of equivalence testing—CDER reviewers' perspective. Proceedings of the Biopharmaceutical Section of American Statistical Association 1999; 214–219.
24. International Conference on Harmonisation. E10: Choice of control group in clinical trials. Federal Register 64(185), September 24, 1999.
25. Julious SA. Sample Sizes for Clinical Trials. 2009 Chapman and Hall/CRC Boca Raton.
26. Garrett AD. Therapeutic equivalence: fallacies and falsification. Statist. Med. 2003; 22: 741–762. DOI: 10.1002/sim.1360.
27. Senn S. Statistical Issues in Drug Development. 1997, Wiley, Chichester.
28. Tu D. On the use of the ratio or the odds ratio of cure rates in therapeutic equivalence clinical trials with binary endpoints. Journal of Biopharmaceutical Statistics 1998;8:135–176.
29. Siegel JP. Equivalence and noninferiority trials. American heart Journal 2000; 139:S166-S170.
30. Kaul S, Diamond GA. Making Sense of Noninferiority: A Clinical and Statistical Perspective on Its Application to Cardiovascular Clinical Trials. Progress in Cardiovascular Diseases 2007; 49(4):284-299.
31. Wang H, Chow SC, Li G. On sample size calculation based on odds ratio in clinical trials. Journal of Biopharmaceutical Statistics 2002; 12(4):471–483.
32. Chow SC, Wang H, Shao J. Sample Size Calculations in Clinical Research, 2nd Edition 2008 Chapman and Hall/CRC Boca Raton.
33. Hintze J. PASS 13 Sample Size Software Chapter 105 Non-Inferiority Tests for One Proportion 2014. ©NCSS, LLC. www.ncss.com
34. Makuch R, Simon R. Sample size requirements for evaluating a conservative therapy. Cancer Treatment Reports 1978; 62(7): 1037-1040
35. Blackwelder WC. "Proving the null hypothesis in clinical trials". Controlled Clinical Trials 1982; 3:345-353.
36. Blackwelder WC, Chang MA. Sample Size for "Proving the Null Hypothesis". Controlled Clinical Trials 1984; 5: 97-105.
37. Heiselbetz C, Edler L. A Sample Size for "Proving the Null Hypothesis". Controlled Clinical Trials 1987; 8: 45-8.
38. Farrington CP, Manning G, Test statistics and sample size formulae for comparative binomial trials with null hypothesis of non-zero risk difference or non-unity relative risk. Statist. Med., 1990; 9: 1447-54.
39. Rodary C, Com-Nogue C, Tournade MF. How to establish equivalence between treatments: a one-sided clinical trial in paediatric oncology, Statist. Med. 1989; 8: 593-8.
40. Miettinen O, Nurminen M. Comparative analysis of two rates, Statist. Med. 1985; 4: 213-26.
41. Hintze J. PASS 13 Sample Size Software Chapter 210 Non-Inferiority Tests for Two Proportions. 2014. ©NCSS, LLC. www.ncss.com.
42. Hintze J. PASS 13 Sample Size Software Chapter 211 Non-Inferiority Tests for the Ratio of Two Proportions. 2014. ©NCSS, LLC. www.ncss.com.
43. Hintze J. PASS 13Sample Size Software Chapter 212 Non-Inferiority Tests for the Odds Ratio of Two Proportions. 2014. ©NCSS, LLC. www.ncss.com.
44. Machin D, Campbell MJ, Fayers PM, Pinol APY. Sample Size Tables for Clinical Studies. 2nd Edition, Oxford, Blackwell Science Ltd. 1997.
45. Machin D, Campbell MJ, Tan SB, Tan SH. Sample Size Tables for Clinical Studies. 3rd Edition, Wiley-Blackwell, 2009.
46. Machin D, Campbell MJ, Tan SB, Tan SH. Sample Sizes for Clinical, Laboratory and Epidemiology Studies. 4th Edition, Oxford, John Wiley & Sons Ltd, 2018.
47. Machin D, Campbell MJ. Sample Size Tables for Clinical Studies. first Edition, Oxford, Blackwell Science Ltd., 1987.
48. Laster LL, Johnson MF. Non-inferiority trials: the 'at least as good as' criterion. Statist. Med., 2003; 22:187–200. DOI: 10.1002/sim.1137.

49. Laster LL, Johnson MF, Kotler ML. Non-inferiority trials: the 'at least as good as' criterion with dichotomous data Statist. Med., 2006; 25:1115–30. DOI: 10.1002/sim.2476.
50. Julious SA, Owen RJ. A Comparison of methods for sample size estimation for non-inferiority studies with binary outcomes. Statistical Methods in Medical Research 2010; 20(6): 595-612.
51. Newcombe RG. interval estimation for the difference between independent proportions: comparison of eleven methods. Statist. Med., 1998; 17: 873-90.
52. Chan ISF. Exact Tests of Equivalence and Efficacy with a Non-zero Lower Bound for Comparative Studies. Stat. Med., 1998, 17, 1403 – 13.
53. Röhmel J, Mansmann U. Unconditional Non-asymptotic One-Sided Tests for Independent Binomial Proportions When the Interest Lies in Showing Non-inferiority and/or Superiority. Biometrical Journal 1999, 2, 149 –70.
54. Chan ISF, Zhang Z. Test-Based Exact Confidence Intervals for the Difference of Two Binomial Proportions. Biometrics, 1999, 55, 1202 – 09.
55. Kang SH, Chen JJ. An approximate unconditional test of non-inferiority between two proportions Statist. Med. 2000;19:2089 -2100.
56. Chan ISF. (2002) Power and Sample Size Determination For Noninferiority Trials Using an Exact Method. Journal of Biopharmaceutical Statistics, 12:4, 457-69, Doi: 10.1081/BIP-120016230.
57. Chan ISF. Proving Non-inferiority or Equivalence of Two Treatments with Dichotomous Endpoints Using Exact Methods. Stat. Methods Med. Res., 2003; 12: 37-58.
58. de Boo TM, Zielhuis GA. Minimization of sample size when comparing two small probabilities in a non-inferiority safety trial Statist. Med., 2004; 23: 1683–99.
59. Hilton JF. Designs of Superiority and Noninferiority Trials for Binary Responses are Noninterchangeable. Biometrical Journal, 2006; 6: 934–47 DOI: 10.1002/bimj.200510288.
60. Hilton JF. Noninferiority trial designs for odds ratios and risk differences Statist. Med. 2010;29: 982—93.

APPENDIX 1

1.1. Machin et al.'s formula

Machin et al.'s formula for calculating sample sizes for non-inferiority studies of the difference between two proportions (5.4, Chapter 4, page 101 [44] and 9.10, Chapter 9, page 109 [45]) needs to be corrected in the case of an unequal allocation ($\phi \neq 1$).

Except for the absolute value of the difference π_1-π_2 at the denominator, the formula:

$$m = \frac{\left(z_{1-\alpha}\sqrt{\phi\bar{\pi}_{1D}\left(1-\bar{\pi}_{1D}\right)+\bar{\pi}_{2D}\left(1-\bar{\pi}_{2D}\right)}+z_{1-\beta}\sqrt{\phi\pi_1\left(1-\pi_1\right)+\pi_2\left(1-\pi_2\right)}\right)^2}{\phi\left(\left|\pi_1-\pi_2\right|-\varepsilon\right)^2}$$

corresponds to Farrington and Manning's formula 4 [38, page 1449]

$$n_1 = \frac{\left(z_{1-\alpha}\sqrt{\tilde{p}_{1D}\left(1-\tilde{p}_{1D}\right)+\frac{\tilde{p}_{2D}\left(1-\tilde{p}_{2D}\right)}{\theta}}+z_{1-\beta}\sqrt{p_1\left(1-p_1\right)+\frac{p_2\left(1-p_2\right)}{\theta}}\right)^2}{\left(p_1-p_2-s_0\right)^2}$$

The term $\theta = n_2/n_1$ of Farrington and Manning's formula 4 [38] is replaced by $\phi = n/m = n_2/n_1$, where n_1 is the sample size of the standard group and n_2 the sample size of the experimental group; in addition, in Farrington and Manning's notation [38], p_1 corresponds to π_1 ($\pi_{\_St}$ in this paper) for the Standard, and p_2 to π_2 ($\pi_{\_Ex}$, in this paper) for the Experimental. Finally, the maximum likelihood estimates in the formulae are indicated with the subscript of a bar [44,45] or a tilde [38]. Eliminating θ at the denominator of the second terms under the square roots in Farrington and Manning's formula [38] means that its corresponding ϕ is the multiplier of the first terms of the square roots at the numerator and the denominator in Machin et al.'s formula.

When the denominators of the two formulae are equal, as can be expected in the case of success probabilities, the numerators are equal and the sample sizes for the Standard drug are the same (n_1 or m,) but, when the calculation is for an unbalanced allocation with $\theta \neq 1$ or $\phi \neq 1$, the results are different. This is because Machin et al.'s formula [44,45] for calculating the coefficients of the cubic equation that gives the maximum likelihood estimates wrongly uses the reciprocal of ϕ (defined as $n/m = n_2/n_1$), as can be seen by the value of the "b" coefficient in Farrington and Manning's equation (b_{FM}) [38]:

$$b_{FM} = -\left(1+\theta+\hat{p}_1+\theta\hat{p}_2+\delta\left(\theta+2\right)\right) = -\frac{n_1+n_2+n_1\hat{p}_1+n_2\hat{p}_2+\delta\left(n_2+2n_1\right)}{n_1}$$

On the contrary, the b coefficient in Machin et al.'s formula (bM) [45,46], with $\phi = n_2/n_1$, is:

$$b_M = -\left(1+\varphi+\varphi\hat{p}_1+\hat{p}_2+\varepsilon\left(1+2\varphi\right)\right)/\varphi = -\left(\frac{n_1+n_2+\hat{p}_1n_2+\hat{p}_2n_1+\varepsilon\left(n_1+2n_2\right)}{n_2}\right)$$

Furthermore, and even more clearly, the "a" coefficients (a_{FM} and a_M) are:

$$a_{FM} = 1+\theta = \frac{n_1+n_2}{n_1} \text{ and } a_M = \frac{1+\varphi}{\varphi} = \frac{n_1+n_2}{n_2}$$

In conclusion, in the case of an unequal allocation and in order to obtain the same results, Machin et al.'s formula [44,45] has to be used with maximum likelihood estimates calculated according to Farrington and Manning [38] or used with the reciprocal of ϕ ($\phi = n_1 / n_2$).

Indeed, in Machin et al.'s equations 5.4 and 5.5 [44, page 101] and equations 9.10 and 9.11 [45, page 109], what needs to be multiplied by the sample size ratio (ϕ) is the standard error of the maximum likelihood estimate of the experimental probability of success ($\bar{\pi}_{2D}$ and π_2,) and not the standard probability of success ($\bar{\pi}_{1D}$ and π_1); finally, ϕ has to be deleted at the denominator. However, the above formulae have been corrected in the last (4th) edition of Machin et al.'s book [46].

1.2. Laster et al.'s approach to failure probabilities

Laster et al.[49] calculated the non-inferiority margin of the difference between two failure probabilities from the non-inferiority margin of the relative risk (and vice versa) by exchanging the role of π_{F_Ex} and π_{F_St} and defining R_{T_F} as π_{F_St} / π_{F_Ex}. The H_0 and H_A of the failure probabilities are therefore formally equal to those of the success probabilities and, consequently, it is necessary to use the same sample size formulae as those used for the success probability. The non-inferiority margin defined by Laster et al. [49] is:

$$R^*_{0\_F} = 1-\frac{\delta^*_{0\_F}}{\pi_{F\_Ex}} \leftrightarrow \delta^*_{0\_F} = \left(1-R^*_{0\_F}\right)\pi_{F\_Ex}$$

a formula that corresponds to that used by us in the case of success probabilities, with π_{S_St} being replaced by π_{S_Ex}. The asterisks at the apex indicate that these quantities are different from those referred to in this paper and are pertinent to only this demonstration.

It should be noted that $R^*_{0\_F}$ depends on π_{F_Ex} which, unlike π_{F_St} or π_{S_St}, does not have only one well-defined value under H_0 and H_A as the values of π_{F_Ex} (or π_{S_Ex}) under H_0 depend on the non-inferiority margin and the true (optimally zero) difference between the standard and experimental probabilities under H_A. This leads to different sample sizes and powers.

However, it is possible to obtain a non-inferiority margin for differences (model 1) that only depends on the known values of π_{F_St} (or π_{S_St}), which are equal under H_0 and H_A.

From the H_0 of the ratio between two probabilities (model 2.1, appropriately called "relative risk" in the case of failure probabilities), we have the following chain of inequalities:

$$H_0 : R^*_{T\_F} \le R^*_{0\_F} \rightarrow \frac{\pi_{F\_St}}{\pi_{F\_Ex}} \le R^*_{0\_F} \rightarrow \pi_{F\_St} \le R^*_{0\_F}\pi_{F\_Ex} \rightarrow -\pi_{F\_Ex} \le -\frac{1}{R^*_{0\_F}}\pi_{F\_St} \rightarrow$$

$$\rightarrow \pi_{F\_St}-\pi_{F\_Ex} \le \pi_{F\_St}-\frac{1}{R^*_{0\_F}}\pi_{F\_St} = \pi_{F\_St}\left(1-\frac{1}{R^*_{0\_F}}\right) \rightarrow$$

$$\rightarrow H_0 : \pi_{F\_St}-\pi_{F\_Ex} \le -\delta_{0\_F} \text{ with } \delta_{0\_F} = \pi_{F\_St}\left(\frac{1-R^*_{0\_F}}{R^*_{0\_F}}\right)$$

This non-inferiority margin corresponds to that shown in Table App.2.3 Failure Probability of the supplementary material,

as $R^*_{O\_F}=1/R_{O\_F}$, it is possible to see that $\delta_{O\_F}=(R_{O\_F}-1)\pi_{F\_St}$ (second row, second column, second formula) by straightforward algebra.

In addition, under H_A, this margin is always larger than that shown by Laster et al. [50], thus leading to lower sample sizes: e.g. according to Laster et al.[50], with $\alpha = 0.05$, $1-\beta = 0.8$, and an equal allocation in the two groups, $\pi_{F\_St} = 0.1$, we obtain:

$R^*_{O\_F}=1/R_{O\_F}$ giving:

$$\delta^*_{O\_F}=\left(1-R^*_{O\_F}\right)\pi_{F\_St}=(1-0.5)0.125=0.0625 \text{ and } n_{F\_St}=876.55\approx 877$$

This result corresponds to that shown between brackets in the first row of Table II of Laster et al.'s paper [49, page 1124].

However, on the basis of our conversion formula, we obtain:

$$\delta_{O\_F}=\pi_{F\_St}\left(\frac{1-R^*_{O\_F}}{R^*_{O\_F}}\right)=0.1\cdot\left(\frac{1-0.5}{0.5}\right)=0.1 \text{ and } n_{F\_St}=219.1373\approx 219$$

In the case of failure probability, Laster et al. change their definition and approach [50, page 1116] adopted for the success probability for which they stated that the non-inferiority margin is "a high percentage or fraction (R_{LB}) of π_{St} ($R_{LB}<1$)". However, this seems to be inconsistent insofar as (R_{LB}) becomes a high percentage or fraction of π_{Ex} (π_{F_Ex}, in our notation) and, consequently, the non-inferiority margin δ_{O_F} of model 1 (D) does not depend on π_{F_St}, but on π_{F_Ex}.

It is also necessary to consider that if, under H_A, $\pi_{F\_St} = \pi_{F\_Ex}$ (as is very sensible), $\delta^*_{O\_F} = \delta^*_{O\_F}$. Consequently, it does not seem to be consistent that the maximum ratio different from 1 under H_A (R_T=0.8) and the maximum non-inferiority margin in terms of a ratio ($R_{LB} = 0.5$) both translate into the very small difference of 0.0625, and it would seem to be more reasonable to obtain our larger difference of 0.1.

Finally, using our approach, it is possible to show that applying the values of the non-inferiority margin obtained directly from a success model to the failure model or vice versa is consistent. This view is also indirectly supported when switching from success model 1 to success model 2.1 and to failure model 2.1 and, finally, to failure model 1. This consistency cannot be demonstrated using Laster et al.'s approach [49] because the settings of success and failure are kept separate.

SUPPLEMENTARY MATERIAL

Appendix 2: Failure probabilities.

2.1 General Methodology.

2.1.1 Formulation of the H_0 and H_A hypotheses

The general methodology is the same as that used for the success probability, except for the formulation of the H_0 and H_A hypotheses, in which the direction of the inequalities is reversed. Using the subscript "F" for failure, these are:

$$H_0 : \theta_F \geq \theta_{0\_F} \text{ for inferiority}$$

$$H_A : \theta_F < \theta_{0\_F} \text{ for non-inferiority}$$

What follows are the differences from the results obtained in the case of success probability.

2.1.2 Statistical significance test

Given the above H_0 formulation, the non-inferiority statistical significance test will always be one-sided (on the left) with the test function given by:

$$\frac{T_F - \mu_{T_F\_H0}}{\sigma_{T_F\_H0}} \approx Z(0;1) \quad (2.1.2)$$

With t_F as the sampling value of T_F and a significance level of α = 0.05 two-sided (or equivalently, 0.025 one-sided), H_0 will be rejected if $z<z_{\alpha/2}$ or $t_F<t_c$, where t_c is the quantile that delimits the critical region. The rejection of the null hypothesis indicates the non-inferiority of the Experimental; however, using the usual approach, the non-inferiority H_0 hypothesis is rejected if the upper limit of the 95% confidence interval is lower than the non-inferiority positive margin.

2.1.3 Sample size calculation

The rationale underlying the sample size calculation is based on the simultaneous occurrence of two events: obtaining a statistically significant result (under H_0) and the rejection of H_0 under H_A:

$$P\{T_F < t_c | H_0\} < \alpha/2 \ \ and \ \ P\{T_F < t_c | H_A\} \geq 1-\beta$$

Solving the above inequalities for t_c gives:

$$\alpha/2 > P(T_F < t_c \,|\, H_0) = P\left(Z < \frac{t_c - \mu_{T_F\_H_0}}{\sigma_{T_F\_H_0}}\right) = \Phi\left(\frac{t_c - \mu_{T_F\_H_0}}{\sigma_{T_F\_H_0}}\right) = P(Z < z_{\alpha/2}) = P(Z < -z_{1-\alpha/2})$$

$$then: \ \frac{t_c - \mu_{T_F\_H_0}}{\sigma_{T_F\_H_0}} \leq -z_{1-\alpha/2} \ \rightarrow \ t_c \leq \mu_{T_F\_H_0} - z_{1-\alpha/2}\sigma_{T_F\_H_0}$$

$$1-\beta \geq P(T_F < t_c \,|\, H_A) = P\left(Z < \frac{t_c - \mu_{T_F\_H_A}}{\sigma_{T\_H_A}}\right) = P(Z < z_{1-\beta})$$

$$then: \frac{t_c - \mu_{T_F\_H_A}}{\sigma_{T_F\_H_A}} \leq z_{1-\beta} = \ \rightarrow \ t_c \leq \mu_{T_F\_H_A} + z_{1-\beta}\sigma_{T_F\_H_A}$$

Finally, by equating the above expressions to t_c, the sample size can be calculated using the following general pivotal formula:

$$z_{1-\alpha/2}\sigma_{T_F\_H_0} + z_{1-\beta}\sigma_{T_F\_H_A} = \mu_{T_F\_H_0} - \mu_{T_F\_H_A} \quad (2.1.3)$$

which has to be explicitly solved for the sample size (n_{S_Ex}) of the Experimental.

The general formula for calculating the sample size of the Experimental is obtained using the algebra shown in paragraph 2.A.3:

$$n_{F\_Ex} = \frac{\left(z_{1-\alpha/2}\sqrt{v^2_{T_F\_H_0}} + z_{1-\beta}\sqrt{v^2_{T_F\_H_A}}\right)^2}{\left(\mu_{T_F\_H_0} - \mu_{T_F\_H_A}\right)^2} \quad (2.1.3.1)$$

where $v^2_{T_F\_H_0} = n_{F\_Ex}\sigma^2_{T_F\_H_0}$; $v^2_{T_F\_H_A} = n_{F\_Ex}\,\sigma^2_{T_F\_H_A}$

The above formula has to be appropriately adapted to the parameters of the considered parameterisations and, in order to allow an unequal allocation, the ratio of the two sample sizes ($k = n_{F\_St} / n_{F\_Ex}$) has to be calculated. The difference in the denominator is inverted but, as it is squared, the result is the same as that obtained from Equation 2.A.3.1

2.1.4 Power of the statistical significance test

The equations are similar to Equations 2.A.4.1 and 2.A.4.2:

$$1-\beta = P\left(T_F \le t_c \mid H_A\right) = P\left(Z \le \frac{t_c - \mu_{T_F\_H_A}}{\sigma_{T_F\_H_A}}\right) = P\left(Z \le \frac{\mu_{T_F\_H_0} + z_{1-\alpha}\sigma_{T_F\_H_0} - \mu_{T_F\_H_A}}{\sigma_{T_F\_H_A}}\right) =$$

$$= P\left(Z \le \frac{z_{1-\alpha/2}\sigma_{T\_H_0} - \left(\mu_{T_F\_H_A} - \mu_{T_F\_H_0}\right)}{\sigma_{T_F\_H_A}}\right) = \Phi\left(\frac{z_{1-\alpha/2}\sigma_{T_F\_H_0} - \left(\mu_{T_F\_H_A} - \mu_{T_F\_H_0}\right)}{\sigma_{T_F\_H_A}}\right) = \quad 2.1.4.1$$

$$= 1 - \Phi\left(\frac{\left(\mu_{T_F\_H_A} - \mu_{T_F\_H_0}\right) - z_{1-\alpha/2}\sigma_{T_F\_H_0}}{\sigma_{T_F\_H_A}}\right)$$

and

$$z_{1-\beta}\sqrt{v^2_{T_F\_H_A}} = \left(\mu_{T_F\_H_A} - \mu_{T_F\_H_0}\right)\sqrt{n_{F\_Ex}} - z_{1-\alpha/2}\sqrt{v^2_{T_F\_H_0}} \rightarrow$$

$$z_{1-\beta} = \frac{\left(\mu_{T_F\_H_A} - \mu_{T_F\_H_0}\right)\sqrt{n_{F\_Ex}} - z_{1-\alpha/2}\sqrt{v^2_{T_F\_H_0}}}{\sqrt{v^2_{T_F\_H_A}}} \quad with \quad 1-\beta = \Phi\left(z_{1-\beta}\right) \quad 2.1.4.2$$

2.2 Models for the Comparison of two Failure proportions

As there are a number of overlaps with the theoretical results shown in section 3A, we shall only consider the differences. The probability of failure of the Standard and Experimental are respectively indicated as π_{F_St} and π_{F_Ex} and, as shown above, the sample probabilities are binomially distributed.

2.2.1 Null (H_0) and alternative (H_A) hypotheses.

Although we maintain the convention of writing the failure probability of the Experimental and Standard in that order, it must be remembered that the inequalities are different. For example, in the case of model 1 (D_F), they are:

$$H_0 : \pi_{F\_Ex} - \pi_{F\_St} \ge \delta_{0\_F}$$
$$H_A : \pi_{F\_Ex} - \pi_{F\_St} < \delta_{0\_F}$$
$$with \;\; 0 < \delta_{0\_F} < 1 - \pi_{F\_St}$$

The non-inferiority margin δ_{0_F} is a value that is appropriate in this setting, and generally fixed at a suitable fraction of π_{F_St}. The other non-inferiority margins also have different limits: $R_{0\_F}>1$ in the case of model 2.1 (R_F), $LR_{0\_F}>0$ in the case of model 2.2 (LR_F), and $OR_{0\_F}>1$ in the case of model 3 (OR_F), which is only considered as $LOR_{0\_F}>0$.

2.2.2 Sampling distribution

As in the case of success probabilities, it is possible to formulate the null and alternative hypotheses, determine sample distributions, and derive the formulae for power and sample sizes for each of the three models.

The sampling distributions of the failure models are the same as those shown in part 3.A, except for the fact that model 1 (D_F) has $-\delta_{0\_F}$ instead of $+\delta_{0\_F}$

2.2.3 Statistical testing

In accordance with the null hypothesis, the statistical tests are one-sided on the left tail of the distribution (instead of being on the right tail as in the case of successes).

2.2.4 Sample size calculation

The formulae for the sample size calculation shown in Table App.2.1 are the same as those obtained in the case of successes, except for the difference model (D_F), which has $-\delta_{0\_F}$ instead of $+\delta_{0\_S}$

2.2.5 Power calculation

Once the sample size has been established as described above, it is once again possible to calculate the power by deriving an *ad hoc* formula as shown in the case of success, or by solving the sample size calculation formula for $z_{1-\beta}$, and then calculating its corresponding probability.

2.2.6 Failure probability tables

Table App.2.1 shows the H_{0_F} hypotheses, sample distributions, and sample size calculation formulae of the three models, with the second model being divided into R_F (model 2.1) and LR_F (model 2.2), and the third model considering OR_F and LOR_F together. The sample size calculation formulae are numbered 3.B.1.4, 3.B.2.1.4, 3.B.2.2.4, and 3.B.3.4 to match the corresponding formulae for the success probability.

Table App.2.2 shows the sample sizes calculated for some values of R_{T_F} (1.18 = 1.00/0.85, 1.11 = 1.00/0.90, 1.05 = 1.00/0.95, and 1.0), and π_{F_St} values ranging from 0.1 to 0.9 at intervals of 0.2, assuming α = 0.025, 1-β = 0.80, and that the non-inferiority margins (expressed as R_{0_F}) are 1.25, 1.15 and 1.05 (which are considered to be suitable for non-inferiority studies) using the three methods of estimating probability.

For example, with π_{F_St} = 0.50, π_{F_EX} = 0.50, giving R_{T_F} = 1.00, δ_{0_F} = 0.125, α = 0.025 and 1-β = 0.80, it is first necessary to calculate $R_{0\_F} = 1 + \delta_{0\_F} / \pi_{F\_St} = 1 + 0.125/0.5 = 1.25$. It is then possible to read that n_{DF_Ex} = 251, n_{RF_Ex} = 322, n_{LRF_Ex} = 315, and n_{LORF_Ex} = 241 in the row with π_{F_St} = 0.5 and R_{0_F} =1.25 and M = 1 (for method 1) in the columns corresponding to R_{T_F} = 1.00. The subsequent two rows show the sample sizes calculated using methods 2 and 3, and it is possible to see that, n_{RF_Ex} = 319 < n_{LRF_Ex} = 326 using method 3. It should be noted that these sample sizes become 809, 1,038, 1,018, and 763 when π_{F_EX} = 0.555 giving R_{T_F} = 1.11.

2.3 Switching non-inferiority margins from one model to another

It is also possible to calculate the pertinent switching formulae for the failure probability (see Table App.2.3) following the same theoretical approach as that used in the case of success probability and starting from their different null hypotheses (H_0); once again it is the standard probability (π_{F_ST}) that plays a pivotal role.

It is worth pointing out that, obtaining the non-inferiority margin of model 2.1 from model 1, we have: $R_{0\_S} = 1 - \frac{\delta_{0\_S}}{\pi_{S\_St}}$ for the success probability, and $R_{0\_F} = 1 + \frac{\delta_{0\_F}}{\pi_{F\_St}}$ for the failure probability.

The formulae for model 1 converted from model 3 are: $\delta_{0\_S} = -\frac{\pi_{S\_St}\left(1-\pi_{S\_St}\right)\left(OR_{0\_S}-1\right)}{1+\pi_{S\_St}\left(OR_{0\_S}-1\right)}$ for the success probability, and $\delta_{0\_F} = \frac{\pi_{F\_St}\left(OR_{0\_F}-1\right)\left(1-\pi_{F\_St}\right)}{1+\pi_{F\_St}\left(OR_{0\_F}-1\right)}$ for the failure probability.

2.3.1 Model 1 (difference/delta: D_F) vs *model 2.1 (ratio: R_F) and Model 2.2 (ln(ratio: LR_F).*

Example 2.3.1. With π_{F_St} = 0.05, π_{F_Ex} = 0.10, and δ_{0_F} = 0.025, it is possible to obtain the equivalent formulation by calculating R_{0_F} = 1 + 0.025 / 0.05 = 1.5 in the case of model 2.1 or by calculating LR_{0_F} = ln(1.5) = 0.405465 for model 2.2.

Alternatively, from π_{F_St} = 0.05 and R_{0_F} = 1.5 or LR_{0_F} = 0.405465, it is straightforward to calculate δ_{0_F} =0.025, by inverting the above relation as: $\delta_{0\_F} = \pi_{F\_St}\,(R_{0\_F} - 1)$ or $\delta_{0\_F} = \pi_{F\_St}\,[\exp(LR_{0\_F})-1]$. The formulated hypotheses of the two models are therefore equivalent.

The same considerations apply in the case of switching from model 1 to model 2.2.

Switching from model 2.1 to model 2.2 only requires changing R_{0_F} to LR_{0_F}.

In addition, it is possible to switch from model 1 to model 2.2 (LR_F) and *vice versa* by using:

$$LR_{0\_F} = ln\left(1+\frac{\delta_{0\_F}}{\pi_{F\_St}}\right) \quad and \quad \delta_{0\_F} = \pi_{F\_St}\left[exp\left(LR_{0\_F}\right)-1\right]$$

2.3.2. Model 1 (D_F) vs *model 3 (odds ratio: OR_F and ln(OR_F)*

Example 2.3.2. With π_{F_St} = 0.05, π_{F_Ex} = 0.10, and δ_{0_F} = 0.025, it is possible to obtain the equivalent formulation for model 3 (OR_F) by calculating:

$OR_{0\_F} = (1 + \delta_{0\_F} / \pi_{F\_St}) \cdot (1 - \pi_{F\_St}) / (1 - \pi_{F\_St} + \delta_{0\_F})$ = 1.54054 and LOR_{0_F} = ln(1.54054) = 0.43213.

Alternatively, from model 3, with π_{F_St} = 0.05, OR_{T_F} = 2.11111, and a non-inferiority margin of OR_{0_F} = 1.54054, it is possible to obtain the equivalent formulation for model 1 (D_F) by calculating:

$$\delta_{0\_F} = -\pi_{F\_St}\,(OR_{0\_F} - 1) \cdot (1 - \pi_{F\_St}) / (1 + \pi_{F\_St} \cdot (OR_{0\_F} - 1)\,) =$$

$$=0.05\cdot(1.54054 - 1) \cdot (1 - 0.05) / (1+0.05\cdot(1.54054 - 1)) = 0.024999 \sim 0.025.$$

It is also possible to calculate:

$\pi_{F\_Ex} = (OR_{T\_F} \cdot \pi_{F\_St}) / (1 + \pi_{F\_St}\cdot(OR_{T\_F} - 1)) = (2.1(1) \cdot 0.05) / (1 + 0.05\cdot(2.1(1) - 1)) = 0.09(9) \sim 0.010$.

The switch from OR_F to LOR_F parameterisation needs no explanation.

2.3.3 Models 2.1 (ratio: R_F) and 2.2 (ln(ratio: LR_F) vs *model 3 (odds ratio: OR_F and ln(OR_F)*

Example 2.3.3.1 With π_{F_St} = 0.05, R_{T_F} = 1.10, and R_{0_F} = 1.20 (Model 2.1), it is possible to obtain the equivalent formulation for model 3 (OR_F) by first calculating:

$$OR_{T\_F} = 1.10\ (1 - 0.05) / (1 - 1.10 \cdot 0.05) = 1.10582$$

and then:

$$OR_{0\_F} = 1.20\ (1 - 0.05) / (1 - 1.20 \cdot 0.05) = 1.21277.$$

LOR_{T_F} and LOR_{0_F} are of course respectively 0.10059 and 0.19290.

The value of OR_{T_F} can also be obtained using the usual formula after having calculated $\pi_{F\_Ex} = R_{T\_F}\cdot\pi_{F\_St}$ = 0.05 · 1.10 = 0.055. Finally, from model 2.2, we first calculate LR_{0_F} = ln(1.20) = 0.18232, and then LOR_{0_F}=0.18232 + ln(0.95) - ln (1 – 1.2 · 0.05) = 0.19290.

The following formulae apply when switching from model 3 to models 2.1 (R_F) and 2.2 (LR_F):

$R_{T\_F} = OR_{T\_F} / (1 + \pi_{F\_St} \cdot (OR_{T\_F} - 1))$ and $R_{0\_F} = OR_{0\_F} / (1 + \pi_{F\_St} \cdot (OR_{0\_F} - 1))$.

Example 2.3.3.2. With π_{F_St} = 0.15, OR_{T_F} = 1.05, and OR_{0_F} = 1.10, it is possible to obtain the equivalent formulation for Model 2.1 by calculating: R_{T_F} = 1.05 / (1 + 0.15·(1.05 - 1)) = 1.04218 and then: R_{0_F} = 1.1 / (1 + 0.15·(1.1 - 1)) = 1.08374.

2.4. Comparison of the sample sizes calculated using the models

In addition to the general considerations made in paragraph 4.A.2 for the success probability, there are also some particular results.

2.4.1. Theoretical results of the sample sizes obtained using all of the models together: 1 (D_F), 2.1 (R_F), 2.2 (LR_F), and 3 (OR_F by the $LOR_F = ln(OR_F)$.

A) Comparison of model 1 (D_F) and models 2.1 (R_F) and 2.2 (LR_F).

Models 2.1 and 2.2 require substantially the same number of patients.

Unlike Laster *et al.* [49], we have found that model 1 (D_F) is always less demanding ($n_{RF\_Ex}>n_{DF\_Ex}$), and sometimes much less demanding, than models 2.1 and 2.2 (see demonstration in Appendix 2.B.1.1).

Example 2.4.1. With α = 0.05 one-sided, power = 1 - β = 0.90, π_{F_St} = 0.2, π_{F_Ex} = 0.2, and δ_{0_F} = 0.05, we obtain: nD_{F_Ex} = 1,099.092 ≈1,100. Alternatively, in the case of model 2.1 and an equivalent non-inferiority margin of R_{0_F}=1+0.05/0.2 = 1.25, we obtain: $n_{R_{F\_Ex}} = 1{,}385.76 \approx 1{,}386$ which corresponds to a 26.0% increase from n_{DF_Ex}; this result is also obtained using maximum likelihood estimates (method 3) according to Farrington and Manning [38]. In particular, in the case of model R_F, the increase in sample size is about 52% if $\pi_{F\_St} = \pi_{F\_Ex}$ = 0.1, about 26% if $\pi_{F\_St} = \pi_{F\_Ex}$ = 0.2, and about 8.8% if $\pi_{F\_St} = \pi_{F\_Ex}$ = 0.6; it then decreases further as π_{F_St} increases, but always remains more demanding than model D_F.

The sample size from model 2.2 (LR_F) is 1,391, which is larger than the 1100 calculated for the D_F model.

B) Asymptotic behaviour study

As in the case of success probabilities, when π_{F_Ex} under H_A tends to its lower limit, which corresponds to $\pi_{S\_St}-\delta_0$ in the case of the D_F model, or when the sample sizes tend to $+\infty$ at a fixed k, non-inferiority margin and π_{F_St}, the following chains of inequalities apply (see Appendix 2.B.1.2).

B.1) $n_{DF\_Ex} < n_{RF\_Ex} = n_{LRF\_Ex}$ and $n_{LORF\_Ex} < n_{RF\_Ex} = n_{LRF\_Ex}$, regardless of the value of π_{F_St};

B.2) $n_{DF\_Ex} \leq n_{LORF\_Ex} < n_{RF\_Ex} = n_{LRF\_Ex}$ when $\pi_{F\_St} \leq (1-\delta_0)/2$, or $n_{LORF\_Ex} < n_{DF\_Ex} < n_{RF\_Ex} < n_{LRF\_Ex}$ when $\pi_{F\_St} > (1-\delta_0)/2$.

The D_F model is therefore less demanding when $\pi_{F\_St} \leq (1-\delta_{0\_F})/2$, and the LOR_F model is less demanding when $\pi_{F\_St} > (1-\delta_{0\_F}) / 2$.

Finally, the R_F and LR_F models are more demanding, but their sample sizes differ by only a few units. For example, with α = 0.05 one-sided, power = 1 - β = 0.90, π_{F_St} = 0.8, π_{F_Ex} = 0.8, and a non-inferiority margin of R_{0_F}=1.0625 (corresponding to δ_{0_F} = 0.05), we obtain: $n_{R_{F\_Ex}} = 1{,}175.02 \approx 1{,}175$ and $n_{LR_{F\_Ex}} = 1{,}179.84 \approx 1{,}180$ using the estimates obtained by means of method 3.

C) Graphical comparisons of the sample sizes obtained using the models

Further results can be obtained by considering the sample size curves calculated for α = 0.05 and 1-β = 0.80, using method 3 at π_{F_St} = 0.3 (Fig. B.1), π_{F_St}=0.5 (Fig. B.2), and π_{F_St}=0.7 (Fig. B.3), with δ_0=0.15 (as a large non-inferiority margin allows a better view of the sample size curves) and π_{F_Ex} ranging over the interval of clinical relevance given by π_{F_St} to $\pi_{F\_St} + \delta_0$.

In the case of π_{F_St} = 0.3, in addition to the fairly parallel pattern of the sample size curves, it is possible to see that: i) the sample sizes of the D_F model are the smallest, with those of the LOR_F model becoming very similar (about 96%) to the values of the D_F model in the case of π_{F_St}= 0.5, and even smaller in the case of π_{F_St}= 0.7; and ii) the sample sizes for the LOR_F model are always smaller than those of the R_F and LR_F models.

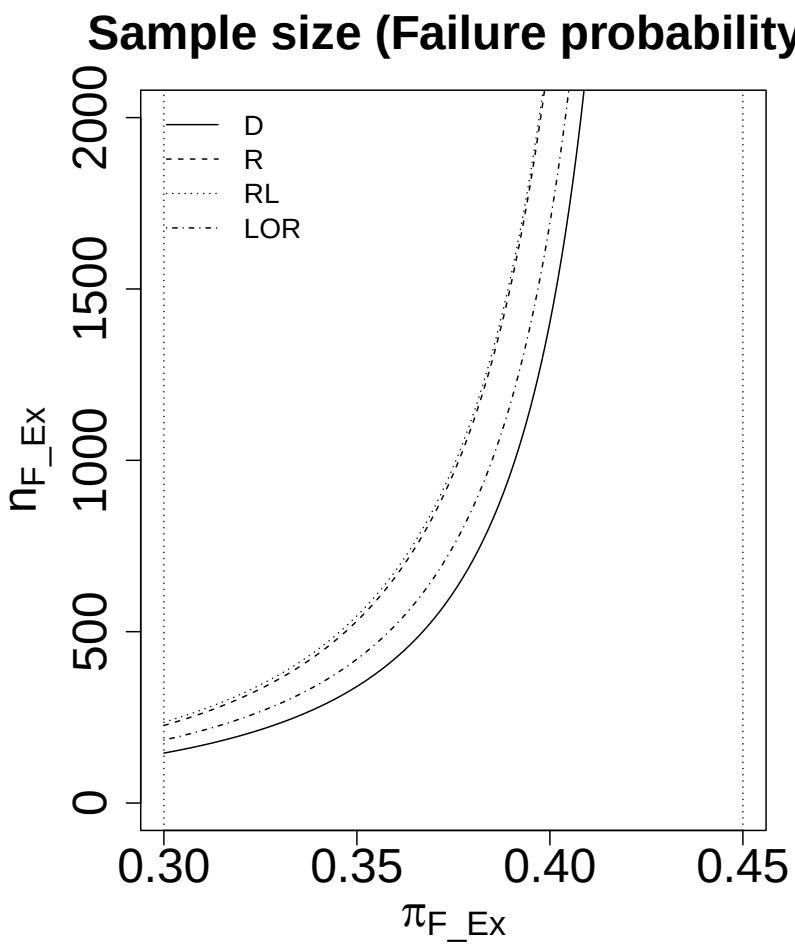


Figure 2.B.4.1: π_{F_St} = 0.3
Sample size curves for π_{S_St} = 0.3 with δ_0=0.15 (because a large non-inferiority margin provides a better vision of the sample size curves of the three models) for α = 0.05, 1-β = 0.80 in function of π_{S_Ex} (ranging from 0.30 to 0.45) using method 3. The curves are for the Difference (D), Ratio (R), Logarithm of the ratio (LR), and the Logarithm of the Odds Ratio (LOR).

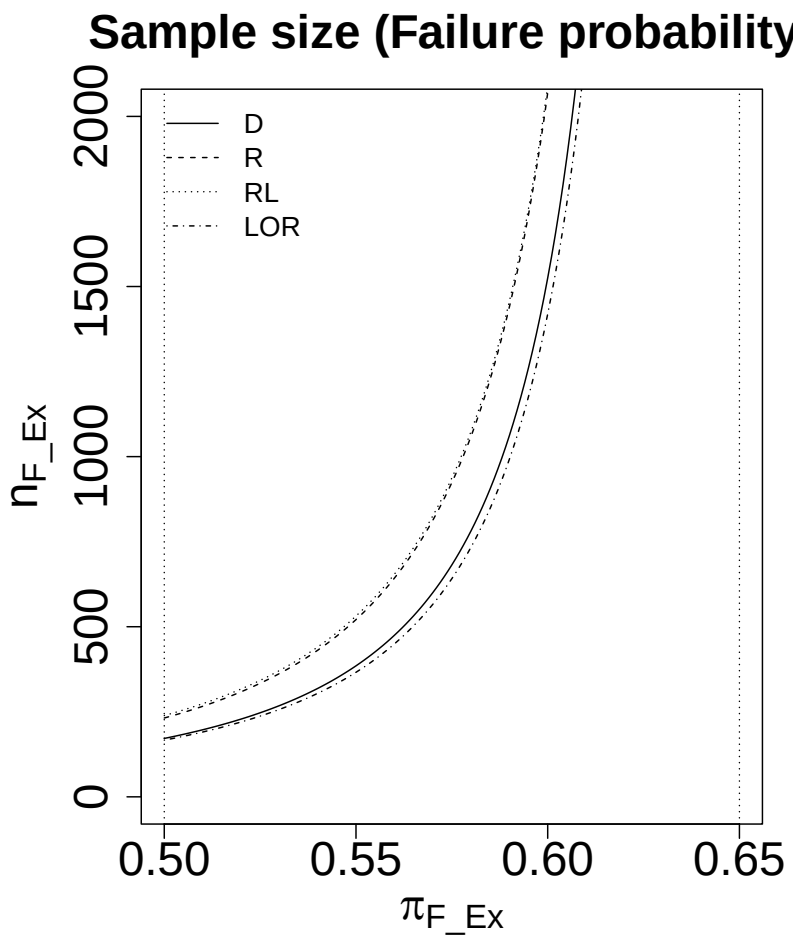


Figure 2.B.4.2: $\pi_{F\_St} = 0.5$

Sample size curves for $\pi_{S\_St} = 0.5$ with δ_0=0.15 (because a large non-inferiority margin provides a better vision of the sample size curves of the three models) for $\alpha = 0.05$, $1\text{-}\beta = 0.80$ in function of π_{S_Ex} (ranging from 0.50 to 0.65) using method 3. The curves are for the Difference (D), Ratio (R), Logarithm of the ratio (LR), and the Logarithm of the Odds Ratio (LOR).

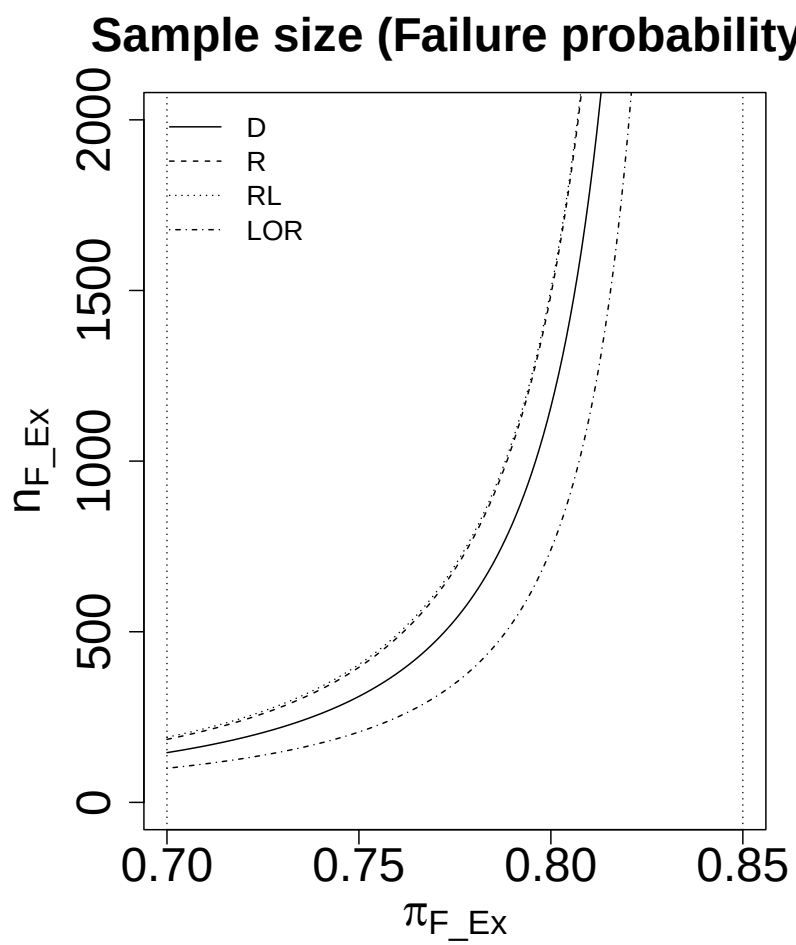


Figure 2.B.4.3: $\pi_{F\_St} = 0.7$

Figure B.3: Sample size curves for $\pi_{S\_St} = 0.7$ with δ_0=0.15 (because a large non-inferiority margin provides a better vision of the sample size curves of the three models) for $\alpha = 0.05$, $1-\beta = 0.80$ in function of π_{S_Ex} (ranging from 0.70 to 0.85) using method 3. The curves are for the Difference (D), Ratio (R), Logarithm of the ratio (LR), and the Logarithm of the Odds Ratio (LOR).

In addition, the pattern of relationships remains substantially the same if the non-inferiority margins are changed; what changes is the entity of the differences in sample sizes.

Finally, changing the methods of estimation does not lead to any evident changes in the relationships except in the case of the R_F and LR_F models, which give sample sizes that differ by only a few units using method 3, are practically equal using method 2, and reverse their relationship with $n_{RLF\_Ex} < n_{RF\_Ex}$ using method 1, but, once again, with differences of only a few units.

Appendix 3

3.A. Success: Asymptotic Behaviour

3.A.1. Model 1 (D_S) *vs* models 2.1 (R_S) and 2.2 (LR_S)

It needs to be remembered that, in the sample size formulae, the numerator of formula 3.A.2.4 for model 2 is less than or (at most) equal to the numerator of formula 3.A.1.4 for model 1, being $R^2_{0\_S} < 1$. Furthermore, the denominators of the two formulae are equal, as is shown by:

$$\left(\pi_{S\_Ex\_H_A} - R_{0\_S}\pi_{S\_St\_H_A}\right)^2 =$$

$$= \left(\pi_{S\_Ex\_H_A} - \left(1 - \frac{\delta_{0\_S}}{\pi_{S\_St\_H_A}}\right)\pi_{S\_St\_H_A}\right)^2 = \left(\pi_{S\_Ex\_H_A} - \pi_{S\_St\_H_A} + \delta_{0\_S}\right)^2$$

The conclusion is that $n_{RS\_Ex} < n_{DS\_Ex}$.

3.A.2. Models 2.1 (R_S) and 2.2 (LR_S)

Considering LR_S as a two-variables function of π_{S_Ex} and π_{S_St}, and applying a first degree Taylor series expansion, starting from $\pi_{S\_Ex\_H0}$ and $\pi_{S\_St\_H0}$, we obtain:

$$LR_S = ln\left(p_{S\_Ex}\right) - ln\left(p_{S\_St}\right) - ln\left(R_{0\_S}\right) \approx$$

$$\approx ln\left(\pi_{S\_Ex\_H_0}\right) - ln\left(\pi_{S\_St\_H_0}\right) - ln\left(R_{0\_S}\right) + \frac{1}{\pi_{S\_Ex\_H_0}}\left(p_{S\_Ex} - \pi_{S\_Ex\_H_0}\right) - \frac{1}{\pi_{S\_St\_H_0}}\left(p_{S\_St} - \pi_{S\_St\_H_0}\right)$$

As $ln\left(\pi_{S\_Ex\_H_0}\right) - ln\left(\pi_{S\_St\_H_0}\right) = ln\left(R_{0\_S}\right)$ and $\pi_{S\_Ex\_H_0} = R_{0\_S}\pi_{S\_St\_H_0}$, we obtain:

$$LR_S \approx \frac{p_{S\_Ex} - \pi_{S\_Ex\_H_0}}{R_{0\_S}\pi_{S\_St\_H_0}} - \frac{p_{S\_St} - \pi_{S\_St\_H_0}}{\pi_{S\_St\_H_0}} = \frac{p_{S\_Ex} - R_{0\_S}p_{S\_St}}{R_{0\_S}\pi_{S\_St\_H_0}} = c \cdot R_S \qquad with\ c = 1/R_{0\_S}\pi_{S\_St\_H_0}$$

and so:

$$LR_S \approx c \cdot R_S \ \ and$$

$$E\left(LR_S\right) \approx c \cdot E\left(R_S\right) = c \cdot \mu_{R_S} = c\left(\pi_{S\_Ex} - R_{0\_S}\pi_{S\_St}\right)$$

$$Var\left(LR_S\right) \approx c^2 Var\left(R_S\right) \ \rightarrow \ v^2_{LR_S} = c^2 v^2_{R_S}$$

In particular:

$$\mu_{LR_{S\_H_0}} \approx c \cdot \mu_{R_{S\_H_0}} = c\left(\pi_{S\_Ex\_H_0} - R_{0\_S}\pi_{S\_St\_H0}\right) = 0; \ \mu_{LR_{S\_H_A}} \approx c \cdot \mu_{R_{S\_H_A}} = c\left(\pi_{S\_Ex\_H_A} - R_{0\_S}\pi_{S\_St\_H_A}\right)$$

$$v^2_{LR_{S\_H_0}} \approx c^2 v^2_{R_{S\_H_0}}; v^2_{LR_{S\_H_A}} \approx c^2 v^2_{R_{S\_H_A}}$$

Including these quantities in the general sample size formula of the LR model, leads to:

$$n_{LR_{S\_Ex}} = \frac{\left(z_{1-\alpha/2}\sqrt{v^2_{LR_{S\_H_0}}} + z_{1-\beta}\sqrt{v^2_{LR_{S\_H_A}}}\right)^2}{\left(\mu_{LR_{S\_H_A}} - \mu_{LR_{S\_H_0}}\right)^2} \approx \frac{c^2\left(z_{1-\alpha/2}\sqrt{v^2_{R_{S\_H_0}}} + z_{1-\beta}\sqrt{v^2_{R_{S\_H_A}}}\right)^2}{\left(c \cdot \mu_{R_{S\_H_A}}\right)^2} =$$

$$= \frac{\left(z_{1-\alpha/2}\sqrt{v^2_{R_{S\_H_0}}} + z_{1-\beta}\sqrt{v^2_{R_{S\_H_A}}}\right)^2}{\left(\mu_{R_{S\_H_A}}\right)^2} = n_{R_{S\_Ex}}$$

The conclusion is therefore that n_{LRS_Ex} is asymptotically equal to n_{RS_Ex} and, in practical terms, the difference is only a few units.

3.A.3. Asymptotic behaviour of the ratio between the sample sizes of the models

A general formula for the sample size calculation of a generic "model T" is:

$$n_{T_{S\_Ex}} = \frac{\left(z_{1-\alpha/2}\sqrt{v^2_{T\_H_0}} + z_{1-\beta}\sqrt{v^2_{T\_H_A}}\right)^2}{\left(\mu_{T\_H_A} - \mu_{T\_H_0}\right)^2} = \frac{\left(z_{1-\alpha/2}\sqrt{v^2_{T\_H_0}} + z_{1-\beta}\sqrt{v^2_{T\_H_A}}\right)^2}{\left(\Delta_T\right)^2}$$

with:

$$v^2_{T\_H_0} = n_{T_{S\_Ex}} \cdot \sigma^2_{T\_H_0} \; ; \; v^2_{T\_H_A} = n_{T_{S\_Ex}} \cdot \sigma^2_{T\_H_A} and \;\; \Delta_T = \mu_{T\_H_A} - \mu_{T\_H_0}$$

and the ratio between the sample sizes of two generic models (T1 and T2) is:

$$\frac{n_{T1_{S\_Ex}}}{n_{T2_{S\_Ex}}} = \frac{\left(z_{1-\alpha/2}\sqrt{v^2_{T1\_H_0}} + z_{1-\beta}\sqrt{v^2_{T1\_H_A}}\right)^2}{\left(z_{1-\alpha/2}\sqrt{v^2_{T2\_H_0}} + z_{1-\beta}\sqrt{v^2_{T2\_H_A}}\right)^2} \cdot \frac{\left(\Delta_{T2}\right)^2}{\left(\Delta_{T1}\right)^2}$$

The asymptotic behaviour of this ratio can be obtained for $H_A \rightarrow H_0$ (μ_{T1_HA} and μ_{T2_HA}, which tend to their non-inferiority limits of respectively μ_{T1_H0} and μ_{T2_H0}), and consequently for $\Delta_{T1} \rightarrow 0$ and $\Delta_{T2} \rightarrow 0$ (herein $\Delta_T \rightarrow 0$):

$$\lim_{\Delta_T \to 0} \frac{n_{T1_{S\_Ex}}}{n_{T2_{S\_Ex}}} = \lim_{\Delta_T \to 0} \frac{\left(z_{1-\alpha/2}\sqrt{v^2_{T1\_H_0}} + z_{1-\beta}\sqrt{v^2_{T1\_H_A}}\right)^2}{\left(z_{1-\alpha/2}\sqrt{v^2_{T2\_H_0}} + z_{1-\beta}\sqrt{v^2_{T2\_H_A}}\right)^2} \cdot \lim_{\Delta_T \to 0} \frac{\left(\Delta_{T2}\right)^2}{\left(\Delta_{T1}\right)^2}$$

$$= \lim_{\Delta_T \to 0} A_{T_1/T_2} \cdot \lim_{\Delta_T \to 0} B_{T_1/T_2}$$

$$with \;\; A_{T_1/T_2} = \frac{\left(z_{1-\alpha/2}\sqrt{v^2_{T1\_H_0}} + z_{1-\beta}\sqrt{v^2_{T1\_H_A}}\right)^2}{\left(z_{1-\alpha/2}\sqrt{v^2_{T2\_H_0}} + z_{1-\beta}\sqrt{v^2_{T2\_H_A}}\right)^2}; \;\; B_{T_1/T_2} = \frac{\left(\Delta_{T2}\right)^2}{\left(\Delta_{T1}\right)^2}$$

Note that, if $\Delta_T \rightarrow 0 \Rightarrow v^2_{T\_H_A} \rightarrow v^2_{T\_H_0}$, then:

$$z_{1-\alpha/2}\sqrt{v^2_{T\_H_0}} + z_{1-\beta}\sqrt{v^2_{T\_H_A}} \rightarrow z_{1-\alpha/2}\sqrt{v^2_{T\_H_0}} + z_{1-\beta}\sqrt{v^2_{T\_H_0}} = \left(z_{1-\alpha/2} + z_{1-\beta}\right)\sqrt{v^2_{T\_H_0}}$$

and so:

$$\lim_{\Delta_T \to 0} A_{T1/T2}\left(\pi_{S\_Ex\_HA}\right) = \frac{\left(z_{1-\alpha/2} + z_{1-\beta}\right)^2 v^2_{T1\_H_0}}{\left(z_{1-\alpha/2} + z_{1-\beta}\right)^2 v^2_{T2\_H_0}} = \frac{v^2_{T1\_H_0}}{v^2_{T2\_H_0}}$$

The above approach can also be applied to $\pi_{S\_Ex\_HA}$ which tends to the non-inferiority limit given by $\pi_{S\_Ex\_H0}$, and to $\pi_{S\_St\_HA} \rightarrow \pi_{S\_St\_H0}$

Section 2.A.3.1 gives an example of the application of this formula to the models LR_S (T_1) and LOR_S (T_2).

3.A.3.1. Calculus of the limits of the ratio between the sample sizes of models LR_S and LOR_S

3.A.3.1.1 Calculus of $\frac{v^2_{LR_S\_H_0}}{v^2_{LOR_S\_H_0}}$ (limit of $A_{T1/T2}$)

$$\frac{v^2_{LR_{S\_H_0}}}{v^2_{LOR_{S\_H_0}}} = \frac{\frac{\left(1-\pi_{S\_Ex\_H_0}\right)}{\pi_{S\_Ex\_H_0}} + \frac{\left(1-\pi_{S\_St\_H_0}\right)}{k\cdot\pi_{S\_St\_H_0}}}{\frac{1}{\pi_{S\_Ex\_H_0}\left(1-\pi_{S\_Ex\_H_0}\right)} + \frac{1}{k\pi_{S\_St\_H_0}\left(1-\pi_{S\_St\_H_0}\right)}} =$$

$$= \frac{\frac{\left(1-R_{0\_S}\pi_{S\_St\_H_0}\right)}{R_{0\_S}\pi_{S\_St\_H_0}} + \frac{\left(1-\pi_{S\_St\_H_0}\right)}{k\pi_{S\_St\_H_0}}}{\frac{1}{R_{0\_S}\pi_{S\_St\_H_0}\left(1-R_{0\_S}\pi_{S\_St\_H_0}\right)} + \frac{1}{k\pi_{S\_St\_H_0}\left(1-\pi_{S\_St\_H_0}\right)}} = \frac{k\left(1-R_{0\_S}\pi_{S\_St\_H_0}\right)+R_{0\_S}\left(1-\pi_{S\_St\_H_0}\right)}{\frac{k}{1-R_{0\_S}\pi_{S\_St\_H_0}} + \frac{R_{0\_S}}{1-\pi_{S\_St\_H_0}}} < 1$$

As the numerator is always less than the denominator, the ratio is <1.

3.A.3.1.2 Calculus of the limit of $B_{LRS/LORS}$: $B_{T_1/T_2} = \frac{\left(\Delta_{T2}\right)^2}{\left(\Delta_{T1}\right)^2}$

Remembering that $\Delta_T \rightarrow 0$ is equivalent to $\pi_{S\_Ex\_HA} \rightarrow \pi_{S\_Ex\_H0} = R_{0\_S}\cdot\pi_{S\_St\_H0}$ (and also to $\pi_{S\_St\_HA} \rightarrow \pi_{S\_St\_H0}$) and applying Hôpital's theorem, we obtain:

$$\lim_{\Delta_T\to 0}\frac{\Delta^2_{LR_S}}{\Delta^2_{LOR_S}} = \lim_{\Delta_T\to 0}\frac{\left[ln\left(R_{T\_S}\right)-ln\left(R_{0\_S}\right)\right]^2}{\left[ln\left(OR_{T\_S}\right)-ln\left(OR_{0\_S}\right)\right]^2} =$$

$$= \lim_{\Delta_T\to 0}\frac{\left[ln\left(\frac{\pi_{S\_Ex\_H_A}}{\pi_{S\_St\_H_A}}\right)-ln\left(R_{0\_S}\right)\right]^2}{\left[ln\left(\frac{\pi_{S\_Ex\_H_A}}{1-\pi_{S\_Ex\_H_A}}\cdot\frac{1-\pi_{S\_St\_H_A}}{\pi_{S\_St\_H_A}}\right)-ln\left(OR_{0\_S}\right)\right]^2} =$$

$$= \lim_{\pi_{S\_Ex\_H_A}\to R_{0\_S}\pi_{S\_St\_H0}}\left(\frac{ln\left(\pi_{S\_Ex\_H_A}\right)-ln\left(\pi_{S\_St\_H_A}\right)-ln\left(R_{0\_S}\right)}{ln\left(\pi_{S\_Ex\_H_A}\right)-ln\left(1-\pi_{S\_Ex\_H_A}\right)+ln\left(\frac{1-\pi_{S\_St\_H_A}}{\pi_{S\_St\_H_A}}\right)-ln\left(OR_{0\_S}\right)}\right)^2 =$$

$$= \lim_{\pi_{S\_Ex\_H_A}\to R_{0\_S}\pi_{S\_St\_H0}}\left(\frac{\frac{1}{\pi_{S\_Ex\_H_A}}}{\frac{1}{\pi_{S\_Ex\_H_A}}+\frac{1}{1-\pi_{S\_Ex\_H_A}}}\right)^2 = \left(\frac{\frac{1}{R_{0\_S}\pi_{S\_St\_H_0}}}{\frac{1}{R_{0\_S}\pi_{S\_St\_H_0}}+\frac{1}{1-R_{0\_S}\pi_{S\_St\_H_0}}}\right)^2 =$$

$$= \left(1-R_{0\_S}\pi_{S\_St\_H_0}\right)^2 < 1$$

3.A.3.1.3. Calculus of the limit of the ratio n_{LRS_Ex}/n_{LORS_Ex}: $Ln_{LRs/LORs}$

$$Ln_{LR_S/LOR_S} = \lim_{\Delta_T \to 0} \frac{n_{LR_{S\_Ex}}}{n_{LOR_{S\_Ex}}} = \frac{v^2_{LR_S\_H_0}}{v^2_{LOR_S\_H_0}} \Big/ \lim_{\Delta_T \to 0} \frac{\Delta_{LR_S}}{\Delta^2_{LOR_S}} = \frac{k\left(1-R_{0\_S}\pi_{S\_St\_H_0}\right)+R_{0\_S}\left(1-\pi_{S\_St\_H_0}\right)}{\frac{k}{1-R_{0\_S}\pi_{S\_St\_H_0}}+\frac{R_{0\_S}}{1-\pi_{S\_St\_H_0}}} \Big/ \left(1-R_{0\_S}\pi_{S\_St\_H_0}\right)^2 =$$

$$= \frac{k+\frac{R_{0\_S}\left(1-\pi_{S\_St\_H_0}\right)}{\left(1-R_{0\_S}\pi_{S\_St\_H_0}\right)}}{k+\frac{R_{0\_S}\left(1-R_{0\_S}\pi_{S\_St\_H_0}\right)}{1-\pi_{S\_St\_H_0}}} < 1$$

and so n_{LR} is asymptotically smaller than n_{LOR}.

3.A.3.2. Limit of the ratio of the sample sizes of models n_{Ds}/n_{Rs} ($Ln_{Ds/Rs}$), n_{Rs}/n_{LRs} ($Ln_{Rs/LRs}$), and n_{Ds}/n_{LORs} ($Ln_{Ds/LORs}$).

Following a similar approach, it is possible to obtain the following results:

$$Ln_{D_S/R_S} = \frac{kR_{0\_S}\left(1-R_{0\_S}\pi_{S\_St\_H_0}\right)+\left(1-\pi_{S\_St\_H_0}\right)}{kR_{0\_S}\left(1-R_{0\_S}\pi_{S\_St\_H_0}\right)+R^2_{0\_S}\left(1-\pi_{S\_St\_H_0}\right)} > 1; \qquad L_{R_S/LR_S} = 1;$$

$$Ln_{LR_S/LOR_S} = \frac{k+\frac{R_{0\_S}\left(1-\pi_{S\_St\_H_0}\right)}{\left(1-R_{0\_S}\pi_{S\_St\_H_0}\right)}}{k+\frac{R_{0\_S}\left(1-R_{0\_S}\pi_{S\_St\_H_0}\right)}{\left(1-\pi_{S\_St\_H_0}\right)}} < 1$$

$$Ln_{D_S/LOR_S} = L_{D_S/R_S} \cdot L_{R_S/LR_S} \cdot L_{LR_S/LOR_S} = \frac{k+\frac{\left(1-\pi_{S\_St\_H_0}\right)}{R_{0\_S}\left(1-R_{0\_S}\pi_{S\_St\_H_0}\right)}}{k+\frac{R_{0\_S}\left(1-R_{0\_S}\pi_{S\_St\_H_0}\right)}{\left(1-\pi_{S\_St\_H_0}\right)}}$$

with $L_{DS/LORS} \geq 1$ if $\pi_{S\_St\_H0} \leq (1+\delta_0)/2$ and $L_{DS/LORS} < 1$ if $\pi_{S\_St\_H0} > (1+\delta_0)/2$.

Conclusions

The ratio limits do not depend on α or β, but only on π_{S_St}, k, and the non-inferiority margin.

Therefore, having fixed k, the non-inferiority margin, and π_{S_St}, the following relations apply:

$L_{DS/RS}>1$, $L_{RS/LRS}= 1$ and $L_{LRS/LORS}<1$, from which it is possible to obtain those relating to sample sizes:

1)- $n_{RS\_Ex} = n_{LRS\_Ex} < n_{DS\_Ex}$ and $n_{RS\_Ex} = n_{LRS\_Ex} < n_{LORS\_Ex}$, regardless of the value of π_{S_St};

2)- $n_{RS\_Ex} = n_{LRS\_Ex} < n_{LORS\_Ex} \leq n_{DS\_Ex}$ when $\pi_{S\_St} \leq (1+\delta_0)/2$, or $n_{RS\_Ex} = n_{LRS\_Ex} < n_{DS\_Ex} < n_{LORS\_Ex}$ when $\pi_{S\_St} > (1+\delta_0)/2$

3.B. Failure: Asymptotic Behaviour

3.B.1. Model 1 (D_F) *vs*. model 2.1 (R_F)

It needs to be remembered that, in the sample size formulae, the numerator of formula 3.B.2.4 for model 2 is greater than or (at most) equal to the numerator of formula 3.B.1.4 for Model 1, as $R^2_{0\_F} > 1$; furthermore, the denominators of the two formulae are equal, as is shown by:

$$\left(\pi_{F\_Ex\_H_A} - R_{0\_F}\pi_{F\_St\_H_A}\right)^2 =$$

$$= \left(\pi_{F\_Ex\_H_A} - \left(1+\frac{\delta_{0\_F}}{\pi_{F\_St\_H_A}}\right)\pi_{F\_St\_H_A}\right)^2 = \left(\pi_{F\_Ex\_H_A} - \pi_{F\_St\_H_A} - \delta_{0\_F}\right)^2$$

The conclusion is that $n_{RF\_Ex} > n_{DF\_Ex}$.

3.B.2. Models 2.1 (R_F) and 2.2 (LR_F)

Following the same approach as that used in section 3.A.2, it is possible to demonstrate that: $n_{LR\_F} \approx n_{R\_F}$.

3.B.3. Asymptotic behaviour of the ratio between the sample sizes of the models.

Following a procedure similar to that used for the success probability (section 3.A.2), the following results are obtained:

$$Ln_{D_F/R_F} = \frac{kR_{0\_F}\left(1-R_{0\_F}\pi_{F\_St\_H_0}\right)+\left(1-\pi_{F\_St\_H_0}\right)}{kR_{0\_F}\left(1-R_{0\_F}\pi_{F\_St\_H_0}\right)+R_{0\_F}^2\left(1-\pi_{F\_St\_H_0}\right)} < 1; \qquad Ln_{R_F/LR_F} = 1;$$

$$Ln_{LR_F/LOR_F} = \frac{k+\dfrac{R_{0\_F}\left(1-\pi_{F\_St\_H_0}\right)}{\left(1-R_{0\_F}\pi_{F\_St\_H_0}\right)}}{k+\dfrac{R_{0\_F}\left(1-R_{0\_F}\pi_{F\_St\_H_0}\right)}{1-\pi_{F\_St\_H_0}}} > 1;$$

$$Ln_{D_F/LOR_F} = \frac{k+\dfrac{\left(1-\pi_{F\_St\_H_0}\right)}{R_{0\_F}\left(1-R_{0\_F}\pi_{F\_St\_H_0}\right)}}{k+\dfrac{R_{0\_F}\left(1-R_{0\_F}\pi_{F\_St\_H_0}\right)}{\left(1-\pi_{F\_St\_H_0}\right)}}$$

with $Ln_{DF/LORF} \leq 1$ if $\pi_{F\_St\_H0} \geq (1-\delta_0)/2$ and $L_{D/LOR} > 1$ if $\pi_{F\_St\_H0} < (1-\delta_0)/2$.

Conclusions

The ratio limits do not depend on α or β, but only on π_{F_St}, k, and the non-inferiority margin.

Then, having fixed k, the non-inferiority margin, and π_{S_St}, the following relations apply:

$L_{DF/RF} > 1$, $L_{RF/LRF} = 1$ and $L_{LRF/LORF} < 1$ from which it is possible to obtain those relating to sample sizes:

1)- $n_{DF\_Ex} < n_{RF\_Ex} = n_{LRF\_Ex}$ and $n_{LORF\_Ex} < n_{RF\_Ex} = n_{LRF\_Ex}$, regardless of the value of π_{F_St};

2)- $n_{DF\_Ex} \leq n_{LORF\_Ex} < n_{RF\_Ex} = n_{LRF\_Ex}$ when $\pi_{F\_St} \leq (1-\delta_0)/2$, or $n_{LORF\_Ex} < n_{DF\_Ex} < n_{RF\_Ex} < n_{LRF\_Ex}$ when $\pi_{F\_St} > (1-\delta_0)/2$

Table App.2.1 Failure Probability. Null Hypothesis (H_0) of the three considered Models (M), together with their sampling distribution, and sample size calculation formulae for the experimental group.

	H_0	Sampling Distribution	$(n)_{F\_Ex} = \ldots (n)_{F\_St} = K \bullet (n)_{F\_Ex}$
M. 1	$\pi_{F\_Ex} - \pi_{F\_St} \geq \delta_{0\_F}$ $with \ \ \delta_{0\_F} > 0$	$D_F = p_{F\_Ex} - p_{F\_St} - \delta_{0\_F}$ $\mu_{D_F} = \pi_{F\_Ex} - \pi_{F\_Ex} - \delta_{0\_F}$ $\sigma^2_{D_F} = \frac{\pi_{F\_St}\left(1-\pi_{F\_St}\right)}{nD_{F\_Ex}} + \frac{\pi_{F\_Ex}\left(1-\pi_{F\_Ex}\right)}{nD_{F\_St}}$	$n_{D_{F\_Ex}} = \frac{\left(z_{1-\alpha/2}\sqrt{\pi_{F\_Ex\_H_0}\left(1-\pi_{F\_Ex\_H_0}\right)+\frac{\pi_{F\_St\_H_0}\left(1-\pi_{F\_St\_H_0}\right)}{k}} + z_{1-\beta}\sqrt{\pi_{F\_Ex\_H_A}\left(1-\pi_{F\_Ex\_H_A}\right)+\frac{\pi_{F\_St\_H_A}\left(1-\pi_{F\_St\_H_A}\right)}{k}}\right)^2}{\left(\pi_{F\_Ex\_H_A} - \pi_{F\_St\_H_A} - \delta_{0\_F}\right)^2}$ 3.B.1.4
M. 2.1	$R_{T\_F} \geq R_{0\_F}\ ; R_{T\_F} = \frac{\pi_{F\_Ex}}{\pi_{F\_St}}$ $R_{0\_F} > 1$ $\left(or: \ \pi_{F\_Ex} - R_0 \pi_{F\_St} \geq 0\right)$	$R_F = p_{F\_Ex} - R_{0\_F} p_{F\_St}$ $\mu_{R_F} = \pi_{F\_Ex} - R_{0\_F}\pi_{F\_St}$ $\sigma^2_{R_F} = \frac{\pi_{F\_Ex}\left(1-\pi_{F\_Ex}\right)}{nR_{F\_Ex}} + R^2_{0\_F}\frac{\pi_{F\_St}\left(1-\pi_{F\_St}\right)}{nR_{F\_St}}$	$n_{R_{F\_Ex}} \frac{\left(z_{1-\alpha/2}\sqrt{\pi_{F\_Ex\_H_0}\left(1-\pi_{F\_Ex\_H_0}\right)+R^2_{0\_F}\frac{\pi_{F\_St\_H_0}\left(1-\pi_{F\_St\_H_0}\right)}{k}} + z_{1-\beta}\sqrt{\pi_{F\_Ex\_H_A}\left(1-\pi_{F\_Ex\_H_A}\right)+R^2_{F\_0}\frac{\pi_{F\_St\_H_A}\left(1-\pi_{F\_St\_H_A}\right)}{k}}\right)^2}{\left(\pi_{F\_Ex\_H_A} - R_{0\_F}\cdot\pi_{F\_St\_H_A}\right)^2}$ 3.B.2.4
M. 2.2	$LR_{T\_F} \geq LR_{0\_F}$ $with \ LR_{T\_F} = \ln\left(R_{T\_F}\right)$ $and \ \ LR_{0\_F} > 0$	$LR_F = ln\left(\frac{p_{F\_Ex}}{p_{F\_St}}\right) - LR_{0\_F}$ $\mu_{LR_F} \approx ln\left(\frac{\pi_{F\_Ex}}{\pi_{F\_St}}\right) - LR_{0\_F} = ln\left(R_{T\_F}\right) - ln\left(R_{0\_F}\right)$ $\sigma^2_{LR_F} \approx \frac{1-\pi_{F\_Ex}}{nLR_{F\_Ex}\cdot\pi_{F\_Ex}} + \frac{1-\pi_{F\_St}}{nLR_{F\_St}\cdot\pi_{F\_St}}$	$n_{LR_{F\_Ex}} \frac{\left[z_{1-\alpha/2}\sqrt{\frac{1-\pi_{F\_Ex\_H_0}}{\pi_{F\_Ex\_H_0}} + \frac{1-\pi_{F\_St\_H_0}}{k\cdot\pi_{F\_St\_H_0}}} + z_{1-\beta}\sqrt{\frac{1-\pi_{F\_Ex\_H_A}}{\pi_{F\_Ex\_H_A}} + \frac{1-\pi_{F\_St\_H_A}}{k\cdot\pi_{F\_StH_A}}}\right]^2}{\left(ln\left(\pi_{F\_Ex\_H_A}/\pi_{F\_St\_H_A}\right) - LR_{0\_F}\right)^2}$ 3.B.2.1.4
M. 3	$OR_{T\_F} \geq OR_{0\_F}$ $with \ OR_{T\_F} = \frac{\pi_{S\_Ex}}{1-\pi_{S\_Ex}} \Big/ \frac{\pi_{S\_St}}{1-\pi_{S\_St}}$ $and \ OR_{0\_F} > 1$ or $LOR_{T\_F} \geq LOR_{0\_F}$ $with \ LOR_{T\_F} = \ln\left(OR_{T\_F}\right)$ $and \ LOR_{0\_F} = \ln\left(OR_{0\_F}\right) > 0$	$LOR_F = \ln\left(OR_F\right) - \ln\left(OR_{0\_F}\right) =$ $= \ln\left(p_{F\_Ex}\right) - \ln\left(q_{F\_Ex}\right) - \ln\left(p_{F\_St}\right)$ $+ \ln\left(q_{F\_St}\right) - LOR_{F\_0}$ $\mu_{LOR_F} = \ln\left(OR_{T\_F}\right) - LOR_{0\_F}$ $\sigma^2_{LOR_F} = \frac{1}{nLOR_{F\_Ex}\cdot\pi_{S\_Ex}\left(1-\pi_{F\_Ex}\right)} +$ $+ \frac{1}{nLOR_{F\_St}\cdot\pi_{S\_St}\left(1-\pi_{F\_St}\right)}$	$n_{LOR_{F\_Ex}} \frac{\left(z_{1-\alpha/2}\sqrt{\frac{1}{\pi_{F\_Ex\_H_0}\left(1-\pi_{F\_Ex\_H_0}\right)} + \frac{1}{k\cdot\pi_{F\_St\_H_0}\left(1-\pi_{F\_St\_H_0}\right)}} + z_{1-\beta}\sqrt{\frac{1}{\pi_{F\_Ex\_H_A}\left(1-\pi_{F\_Ex\_H_A}\right)} + \frac{1}{k\cdot\pi_{F\_St\_H_A}\left(1-\pi_{F\_St\_H_A}\right)}}\right)^2}{\left(ln\left(\pi_{F\_Ex\_H_A}\right) - ln\left(1-\pi_{F\_Ex\_H_A}\right) - ln\left(\pi_{F\_St\_H_A}\right) + ln\left(1-\pi_{F\_St\_H_A}\right) - LOR_{0\_F}\right)^2}$ 3.B.3.4

The sample size calculation formulae are numbered as: 3.B.1.4, 3.B.2.4, 3.B.2.1.4, and 3.B.3.4 for consistency with the corresponding for the success probability.

Table App.2.2 Failure Probability. Sample sizes for $\alpha = 0.025$ and $1-\beta = 0.80$

			R_{T_F}=1.18*	1.18*	1.18*	1.18*	1.11$	1.11$	1.11$	1.11$	1.05#	1.05#	1.05#	1.05#	1.00	1.00	1.00	1.00
π_{F_St}	R_{0_F}	M	D_F	R_F	LR_F	LOR_F	D_F	R_F	LR_F	LOR_F	D_F	R_F	LR_F	LOR_F	D_F	R_F	LR_F	LOR_F
0.1	1.25	1	28,135	35,485	35,237	34,206	7,681	9,740	9,618	9,348	3,711	4,731	4,651	4,524	2,260	2,896	2,837	2,762
0.1		2	28,119	35,321	35,398	34,333	7,673	9,655	9,691	9,406	3,707	4,671	4,694	4,559	2,258	2,849	2,865	2,784
0.1		3	28,178	35,312	35,409	34,333	7,706	9,650	9,697	9,406	3,730	4,668	4,698	4,559	2,277	2,847	2,868	2,784
0.1	1.15	1	-	-	-	-	97,967	11,3030	11,2747	11,0798	15,248	17,651	17,550	17,257	6,279	7,292	7,233	7,115
0.1		2	-	-	-	-	97,948	11,2844	11,2932	11,0946	15,243	17,578	17,607	17,304	6,276	7,245	7,260	7,137
0.1		3	-	-	-	-	98,018	11,2834	11,2945	11,0946	15,271	17,574	17,612	17,304	6,296	7,242	7,263	7,137
0.1	1.05	1	-	-	-	-	-	-	-	-	-	-	-	-	56,512	59,408	59,349	59,022
0.1		2	-	-	-	-	-	-	-	-	-	-	-	-	56,509	59,361	59,377	59,045
0.1		3	-	-	-	-	-	-	-	-	-	-	-	-	56,529	59,359	59,380	59,045
0.3	1.25	1	7,071	8,977	8,898	7,898	1,954	2,488	2,452	2,187	954	1,218	1,196	1,071	586	751	736	660
0.3		2	7,055	8,937	8,952	7,924	1,947	2,468	2,476	2,199	950	1,205	1,210	1,078	583	741	745	665
0.3		3	7,060	8,932	8,967	7,924	1,950	2,465	2,484	2,199	952	1,203	1,215	1,078	585	739	749	665
0.3	1.15	1	-	-	-	-	24,924	28,830	28,739	26,826	3,919	4,542	4,513	4,224	1,628	1,890	1,875	1,758
0.3		2	-	-	-	-	24,906	28,784	28,801	26,857	3,914	4,525	4,532	4,233	1,625	1,880	1,884	1,763
0.3		3	-	-	-	-	24,912	28,778	28,818	26,857	3,916	4,523	4,538	4,233	1,627	1,879	1,888	1,763
0.3	1.05	1	-	-	-	-	-	-	-	-	-	-	-	-	14,651	15,402	15,387	15,060
0.3		2	-	-	-	-	-	-	-	-	-	-	-	-	14,648	15,392	15,396	15,065
0.3		3	-	-	-	-	-	-	-	-	-	-	-	-	14,651	15,390	15,400	15,065
0.5	1.25	1	2,858	3,675	3,630	2,685	809	1,038	1,018	763	402	516	505	382	251	322	315	241
0.5		2	2,842	3,660	3,663	2,701	802	1,031	1,033	771	398	512	514	387	248	319	321	243
0.5		3	2,843	3,662	3,684	2,701	802	1,031	1,044	771	398	512	521	387	248	319	326	243
0.5	1.15	1	-	-	-	-	10,316	11,989	11,938	10,089	1,653	1,920	1,905	1,622	698	810	804	687
0.5		2	-	-	-	-	10,297	11,972	11,975	10,107	1,648	1,915	1,917	1,627	695	807	809	690
0.5		3	-	-	-	-	10,297	11,973	11,998	10,107	1,648	1,915	1,926	1,627	695	807	815	690
0.5	1.05	1	-	-	-	-	-	-	-	-	-	-	-	-	6,279	6,601	6,594	6,269
0.5		2	-	-	-	-	-	-	-	-	-	-	-	-	6,276	6,598	6,600	6,271
0.5		3	-	-	-	-	-	-	-	-	-	-	-	-	6,276	6,598	6,605	6,271
0.7	1.25	1	1,053	1,403	1,373	556	318	416	404	172	166	215	209	93	108	138	135	62
0.7		2	1,037	1,399	1,396	590	311	415	415	186	162	215	215	101	105	138	139	67
0.7		3	1,051	1,418	1,435	590	316	423	433	186	165	219	226	101	107	141	147	67
0.7	1.15	1	-	-	-	-	4,055	4,772	4,737	3,038	682	797	788	517	299	347	344	230
0.7		2	-	-	-	-	4,036	4,767	4,764	3,074	677	796	796	527	296	348	348	235
0.7		3	-	-	-	-	4,049	4,786	4,805	3,074	681	802	811	527	298	351	357	235
0.7	1.05	1	-	-	-	-	-	-	-	-	-	-	-	-	2,691	2,829	2,826	2,502
0.7		2	-	-	-	-	-	-	-	-	-	-	-	-	2,688	2,829	2,830	2,507
0.7		3	-	-	-	-	-	-	-	-	-	-	-	-	2,690	2,833	2,839	2,507
0.9	1.25	1	-	-	-	-	45	71	63	-	35	47	44	-	28	36	35	-
0.9		2	-	-	-	-	-	-	-	-	-	-	-	-	-	38	38	-
0.9		3	-	-	-	-	73	107	115	-	48	67	73	-	36	48	54	-
0.9	1.15	1	-	-	-	-	577	763	737	-	143	173	167	-	78	90	89	-
0.9		2	-	-	-	-	-	-	-	-	137	175	174	-	75	92	92	-
0.9		3	-	-	-	-	694	899	913	-	168	207	214	-	89	107	112	-
0.9	1.05	1	-	-	-	-	-	-	-	-	-	-	-	-	698	733	733	417
0.9		2	-	-	-	-	-	-	-	-	-	-	-	-	695	736	736	440
0.9		3	-	-	-	-	-	-	-	-	-	-	-	-	713	756	762	440

Legend: π_{F_St} = true failure probability for the Standard drug, R_{0_F} = non-inferiority margin expressed in the ratio scale, M = Method 1, 2, and 3 (see text); R_{T_F} = true ratio between the true success probability for the Experimental drug (π_{F_Ex}) and π_{F_St}; D_F = Difference, R_F = Ratio, LR_F = ln(R_F), LOR_F = ln (Odds Ratio). 1.18* = 1.00/0.85;1.11$ = 1.00/0.90; 1.05# = 1.00/0.95. The "–" sign means that it is a case incompatible with non-inferiority and the "." sign means that the denominator of the sample size formula is equal to 0 (sample size tends to infinity)

Table App.2.3. Failure Probability. Formulae for switching from a model to another of the three considered models

From the Model:	To the Model:		
Model 1	Model 2.1	Model 2.2	Model 3
$\begin{cases}\pi_{F\_Ex}\\ \pi_{F\_St}\\ \delta_{0\_F}\end{cases}\rightarrow$	$R_{T\_F}=\dfrac{\pi_{F\_Ex}}{\pi_{F\_St}}$ $R_{0\_F}=1+\dfrac{\delta_{0\_F}}{\pi_{F\_St}}$	$LR_{T\_F}=ln\left(\dfrac{\pi_{F\_Ex}}{\pi_{F\_St}}\right)$ $LR_{0\_F}=ln\left(1+\dfrac{\delta_{0\_F}}{\pi_{F\_St}}\right)$	$OR_{T\_F}=\dfrac{\pi_{F\_Ex}}{\pi_{F\_St}}\cdot\dfrac{1-\pi_{F\_St}}{1-\pi_{F\_Ex}}$ $OR_{0\_F}=\left(1+\dfrac{\delta_{0\_F}}{\pi_{F\_St}}\right)\left(\dfrac{1-\pi_{F\_St}}{1-\left(\pi_{F\_St}+\delta_{0\_F}\right)}\right)$ $LOR_{T\_F}=ln\left(OR_{T\_F}\right);\ LOR_{0\_F}=ln\left(OR_{0\_F}\right)$
Model 2.1	Model 1	Model 2.2	Model 3
$\begin{cases}\pi_{F\_St}\\ R_{T\_F}\\ R_{0\_F}\end{cases}\rightarrow$	$\pi_{F\_Ex}=R_{T\_F}\cdot\pi_{F\_St}$ $\delta_{0\_F}=\left(R_{0\_F}-1\right)\pi_{F\_St}$	$LR_{T\_F}=ln\left(R_{T\_F}\right)$ $LR_{0\_F}=ln\left(R_{0\_F}\right)$	$OR_{T\_F}=R_T\dfrac{1-\pi_{F\_St}}{1-R_{T\_F}\cdot\pi_{F\_St}}$ $OR_{0\_F}=R_0\dfrac{1-\pi_{F\_St}}{1-R_{0\_F}\cdot\pi_{F\_St}}$ $LOR_{T\_F}=ln\left(OR_{T\_F}\right);\ LOR_{0\_F}=ln\left(OR_{0\_F}\right)$
Model 2.2	Model 1	Model 2.1	Model 3
$\begin{cases}\pi_{F\_St}\\ LR_{T\_F}\\ LR_{0\_F}\end{cases}\rightarrow$	$\pi_{F\_Ex}=exp\left(LR_{T\_F}\right)\cdot\pi_{F\_St}$ $\delta_{0\_F}=\left[exp\left(LR_{0\_F}\right)-1\right]\cdot\pi_{F\_St}$	$R_{T\_F}=exp\left(LR_{T\_F}\right)$ $R_{0\_F}=exp\left(LR_{0\_F}\right)$	$OR_{T\_F}=exp\left(LR_{T\_F}\right)\cdot\dfrac{1-\pi_{F\_St}}{1-exp\left(LR_{T\_F}\right)\cdot\pi_{F\_St}}$ $OR_{0\_F}=exp\left(LR_{0\_F}\right)\dfrac{1-\pi_{F\_St}}{1-exp\left(LR_{0\_F}\right)\cdot\pi_{F\_St}}$ $LOR_{T\_F}=ln\left(OR_{T\_F}\right);\ LOR_{0\_F}=ln\left(OR_{0\_F}\right)$
Model 3	Model 1	Model 2.1	Model 2.2
$\begin{cases}\pi_{F\_St}\\ OR_{T\_F}\\ OR_{0\_F}\end{cases}\rightarrow$	$\pi_{S\_Ex}=\dfrac{OR_{T\_F}\cdot\pi_{F\_St}}{1+\pi_{F\_St}\left(OR_{T\_F}-1\right)}$ $\delta_{0\_F}=\dfrac{\pi_{F\_St}\left(OR_{0\_F}-1\right)\left(1-\pi_{F\_St}\right)}{1+\pi_{F\_St}\left(OR_{0\_F}-1\right)}$	$R_{T\_F}=\dfrac{OR_{T\_F}}{1+\pi_{F\_St}\left(OR_{T\_F}-1\right)}$ $R_{0\_F}=\dfrac{OR_{0\_F}}{1+\pi_{F\_St}\left(OR_{0\_F}-1\right)}$	$LR_{T\_F}=ln\left(OR_{T\_F}\right)-ln\left(1+\pi_{F\_St}\left(OR_{T\_F}-1\right)\right)$ $LR_{0\_F}=ln\left(OR_{0\_F}\right)-ln\left(1+\pi_{F\_St}\left(OR_{0\_F}-1\right)\right)$