Avenue E. Mounier 83/11
1200 Brussels
Belgium
Tel: +32 2 774 1611
Email: eortc@eortc.be
www.eortc.org

Design of Randomization and Stratification of Clinical Trials

ST-002-WIN-02
Version: 2
ALWAYS REFER TO INTRANET TO CHECK THE VALIDITY OF THIS DOCUMENT

Author *Biostatistician* Corneel Coens	Signature:	Date: 28/7/15
Approved and authorized by *Methodology Director* Jan Bogaerts	Signature:	Date: 30.7.2015

1 PURPOSE

To describe the design of randomization parameters in trials where EORTC handles the randomization.

2 INSTRUCTION

2.1 General Design

The intention of a randomization is to assign treatment arm allocations in such a way as to produce comparable groups in terms of general patient characteristics and specific key factors that affect the probable course the disease would take, while minimizing potential selection bias by ensuring a degree of unpredictability in the treatment allocation. [Ref 1, 2].

The Statistician carefully considers the following elements during the design of a study, and describes these in the statistical chapter of the protocol.

- Timing of the randomization: in order to minimize non-compliance and differential effects between randomized arms, the randomization has to be performed as close as possible prior to the start of the defined protocol interventions. In some trials, the setup is such that the assigned treatment arm does not take effect in all randomized patients. In such cases, a temporary blinding (until start of treatment) can be considered.
- Randomization ratio: the recommendation is to use a 1:1 ratio (i.e. placing approximately equal numbers of patients on each treatment arm) across all treatment arms as this is most efficient in terms of total patient numbers for comparisons. Deviations can be of interest in some situations, but need to be explicitly justified. For example, with a 2:1 randomization, there is an increase of 12.5 percent in the required number of events. Also, estimates of efficacy and safety in the smaller arm(s) are less precise. It might be of interest if already sufficient information is known about the control arm and more patients are needed on the experimental arm for adequate safety documentation or if the 2:1 randomization would make the trial more feasible by increasing the patient appeal. In a randomized trial where there is no formal comparison (eg. randomized phase II trial based on single arm design in the experimental arm) the sample size should be justified for each arm and the randomization ratio adjusted accordingly.
- Method of randomization: two methods are available for implementation in ORTA (On-line Randomized Trial Access): a dynamic minimization algorithm and a static permuted block method. The recommended method is the minimization method as it is more flexible and allows more levels of stratification to be included than the permuted blocks. Since the minimization method randomizes each new patient based on the characteristics of the already randomized patients, it is less susceptible to deviations in assumptions or timing of the analysis. The permuted block randomization can be used when justifiable, such as requirement from external partner or logistical constraints. The two methods are described in more detail in the next paragraph. Hybrid methods (combining dynamic and static designs, whereby a permuted block list is used but allocations may be skipped based on imbalance scores) or adaptive randomization (where randomization ratios depend on observed patient outcomes) are not available through ORTA.
    - Stratification factors: **The randomization should be stratified for a limited number of important prognostic factors applicable to the study population**. The stratification factors with their respective distinct levels should be listed in the protocol. In case institution is a stratification factor, this should be stated explicitly but the individual institutions do not need to be specified. If there is justifiable concern that the listing of the stratification levels would allow prediction of the allocation, the listing can be documented in a separate document. Care should be taken that the levels of each

stratification factor are relevant, reliable and obtainable at the time of the randomization.
More recommendations about the choice of stratification factors are given per method in the next paragraph.

- If more than one randomization is planned within a single trial, the following must be considered:
  - Later randomizations have to be stratified for the treatment allocation of the earlier one(s).
  - For 2x2 factorial designs, where both randomizations occur at the same moment, a single randomization over the possible treatment combinations is advised (ie. for a classical 2x2, a single 4 arm randomization would be employed).

It is generally recommended to perform simulations to assess the performance of the chosen stratification design. The performance should assess both the predictability of the allocations and the balance of the stratification factors over the treatment arms at the moment(s) of analysis.

2.2 Minimization Algorithm

The algorithm used at EORTC for dynamic treatment allocation is a modified version of the methods described in Pocock and Simon (Ref 3) and Freedman and White (Ref 4, 5): the variance method with equal weights ($w_i = 1$) and equal assignment probabilities $^1/_N$ where N is the number of treatments. The algorithm assigns a treatment to patients that are entered sequentially into a clinical trial by choosing among the pre-defined study treatments in such a way that for each of the different stratification factor levels, the number of patients assigned to each treatment group is approximately the same.

When a new patient with known levels for all of the stratification factors is entered in the trial, the algorithm evaluates the degree of imbalance across treatments for all patients already randomized that present the same characteristics in terms of each of the stratification factors separately. This is done in order to split the possible treatment assignment choices into 2 mutually exclusive sets, namely those treatment choices that result in a total imbalance change (from the minimum) lower or equal than a preset threshold factor (set A) and those that will not (set B).

The total imbalance for a certain treatment is calculated as follows. For each factor, the number of existing patients with the same level as that of the patient to be randomized is counted in the different treatment arms. The count is repeated for all factors and the total sum across all factors is then determined per treatment arm. These totals are referred to as the "total imbalance score" for a certain treatment. Next, the difference between each treatment group total and the lowest total is determined for each treatment. This difference is the quantity that is compared to the threshold.

Once the set of potential treatment arms (set A) has been determined, the treatment allocation is based on assigning to each treatment in this set a random number that is sampled from a uniform distribution in the range [0, 1). The compiler's random number generator generates this number (Visual Basic RND function uses linear-congruential method for pseudo-random number generation). Once a number has been assigned to all predefined potential treatments, then a treatment is allocated by choosing the treatment that has the maximum random number assigned to it. It should be noted that this is equivalent to randomly choosing among the available choices, each having the same probability of being assigned.

2.2.1 Random component (P)

Based on simulations, the default settings (most notably the inclusion of institution as stratification factor and the threshold equal to the number of stratification factors) result in a sufficient proportion of random (i.e. not deterministic) assignments (typically 60-70% in a 2-arm trial). If desirable, a random selection probability

(P) can be added to the algorithm. If this random component is added, the algorithm chooses a treatment arm at random (ie. with equal probability) among all treatment arms (ie. A U B) with probability P. The treatment arm is determined on basis of the above-described minimization algorithm (ie. at random in A only) with probability 1-P. This is achieved by generating a random number that is sampled from a uniform distribution in the interval [0, 1] which is then compared to the value of P. The value of this random component P needs to be specified before the implementation of the randomization algorithm and should be set to a value < 0.25, preferably in the range of 0.15 to 0.05. In general, it is recommend to run simulations for the foreseen setup of the trial to document the consequences of particular choices of P, in terms of distribution of prognostic factors and predictability. By default the random component is set equal to 0, resulting in always selecting an arm from set A. By doing so the treatment choice can still result in a total imbalance score that is larger than the minimum but the difference cannot be larger than the preset threshold.

2.2.2 Threshold

The particular threshold value used for splitting the treatment choices into two sets needs to be specified before the implementation of the randomization algorithm. It is by default set to be the total number of stratification factors but can be set to any other value if needed. In the default case, if the difference in imbalance for a treatment arm is greater than the total number of stratification factors then that treatment falls into set B and set A otherwise. Thus the maximum degree of imbalance is then bounded by the number of patients in the particular strata and the number of stratification factors + 1.

2.2.3 Randomization ratio

The algorithm is limited to a maximum number of 99 different treatment arms by technical constraints. Randomization occurs by default so that the total number of patients is approximately equally spread over the available treatments arms (example: a 2-arm randomization will by default be a 1:1 randomization).

If different proportions per treatment arm are required, additional treatment arms need to be defined in order to reach the desired ratio (example: a 2-arm randomization 2:1 ratio can be achieved by defining three treatment arms where two arms will yield the same allocation label). The ratio needs to be specified before the implementation of the randomization algorithm and needs to be listed in the study outline and protocol.

2.2.4 Design considerations

- By default, institution is considered a stratification factor but it can be removed (or replaced) when justified. Institution is often a strong prognostic factor and should be kept as stratification factor if feasible. In order to ensure the unpredictability of the treatment assigned, it is recommended that all trials be stratified by institution and at least one additional factor.
- While several publications suggest that minimization can efficiently handle a large number of stratification factors, it is recommended to limit the number of factors so that the resulting number of strata does not exceed the number of patients per treatment arm [Ref 6, 7, 8]. In addition, as the resulting analysis needs to be stratified for the non-institutional factors, the number of events per strata should be adequate in order to prevent substantial loss of power [Ref 9]. ***Therefore the recommendation is that the number of stratification factors is bounded so that the total number of all strata (including institution as a factor) does not exceed the number of patients per treatment arm and the total number of non-institutional strata times 5 does not exceed the number of events per treatment arm.***

- If a prognostic interaction between two or more stratification factors is expected, then a combined summary factor should be used at the stratification level rather than these individual factors.
- Factor levels with expected small numbers of subjects must be considered carefully if not to be merged with other levels.

It is generally recommended to perform simulations to assess the performance of the chosen stratification design via the %DYNBALSIM macro.

If the predictability is too high, this can be remedied by:

- Increasing the value of the random component P (by default is set to 0).
- Increasing the value of the threshold parameter.

If strict balance over the treatment arms is an issue then:

- The threshold level can be decreased.
- Institution can be removed or replaced by a factor with fewer levels (eg. country).
- Where relevant, two or more stratification factors can be combined into a simpler summary factor.
- An extra stratification factor with only one level (ie. equal for all patients) can be added [Ref 9].

2.2.5 Technical document

Two documents detailing the minimization algorithm, one in descriptive manner and one in more formal detail, are available from the J:\UNIT\Stat\3. Minimization\SOPs & descriptive documents directory for external circulation if needed

2.3 Permuted Block Algorithm

The algorithm used at EORTC for static treatment allocation is the permuted block design [Ref 10, 11]. The algorithm generates an allocation list for the whole trial prior to the first patient randomization and assigns treatment allocations accordingly. For each distinct strata (ie. unique level of all stratification factors), a series of blocks of length *l* are created where *l* is multiple of the number of treatment arms or of the sum of the allocation ratios in case of unequal treatment ratios. Each block contains an ordered permuted series of treatment allocations such that the required treatment balance is reached within the blocks. Whenever a new patient is entered in the trial, the appropriate strata are identified and the first available treatment allocation is assigned to that patient. Unpredictability is driven by the random permutation within each block. Imbalance between the two treatment arms will be due to incomplete blocks only.

2.3.1 Randomization ratio

If different proportions per treatment arm are required, the ratio can be specified under the form of $R_1{:}R_2{:}\ldots R_m$ where m = the total number of treatment arms and R_i is the relative ratio of treatment arm i. The values of R_i are limited to integers only. It should be noted that permissible block lengths need to be a multiple of $\sum R_i$ (i = 1 to m). Therefore the choice of the R_i will impact on the possible block length which in turn will impact on the achievable balance. The ratios need to be specified before the implementation of the randomization algorithm and need to be listed in the study outline and protocol.

2.3.2 Block length

The block length needs to be multiple of $\sum R_i$ (i = 1 to m) where $R_1{:}R_2{:}\ldots R_m$ are the treatment allocation ratios and m is the total number of treatment arms. The ORTA system allows the specification of multiple block lengths with relative frequency. The resulting allocation list will generate then blocks where for each block the length will be chosen at random from the specified block lengths with a probability relative to the designated frequency value.

If there is justifiable concern that the listing of the block length(s) would allow prediction of the allocation, the listing can be documented in a separate document.

2.3.3 Design considerations

- The achieved balance is mainly limited by the number of incomplete blocks. The resulting treatment imbalance will be bounded by the maximum imbalance within each block multiplied by the number of strata. The maximum imbalance within a block increases with increasing block length.
- Unpredictability is driven mainly by knowledge of the stratification levels and the block length. Longer block lengths will lead to fewer deterministic allocations.
- It is recommended to limit the number of strata by careful selection of the factors and the corresponding levels. The resulting strata and block length should allow for at least 1 complete block per strata [Ref 12]. In addition, as the resulting analysis needs to be stratified, the number of events per strata should be adequate in order to prevent substantial loss of power [Ref 9]. In case of equally distributed strata (ie. the expected number of patients enrolled is equal for each strata), this implies that the number of strata multiplied by the block length should not exceed the sample size. In case of unequally distributed strata, the expected number of patients in the smallest stratum should be equal to at least the block length. Strata with expected small numbers of subjects must be considered carefully if not to be merged.
- It is recommended to choose a varying block length rather than a single fixed value. Compared to a fixed value l_1, a block length alternating between l_2 and l_1 will decrease the predictability of the allocations and improve the balance if $l_2 < l_1$.

It is generally recommended to perform simulations to assess the performance (predictability of the allocations and the balance of the stratification factors over the treatment arms at the moment(s) of analysis) of the chosen stratification design via PROC PLAN in SAS.

<u>If the predictability is too high, this can be remedied by:</u>

- Choosing a longer block length or opting for a varying block length in case of a single block length value.
- Adding additional permissible block lengths in case of a varying block length design.
- Not listing the stratification levels and/or block lengths in any communication to the participating investigators.

<u>If strict balance over the treatment arms is an issue then:</u>

- The block length should be decreased. In case of varying block lengths, a higher frequency rate can be given to the shorter block length.
- The number of strata should be decreased. Note that combining two or more stratification factors into their factorial summary factor will not decrease the number of resulting strata.

2.4 Implementation

The procedure for the implementation of the ORTA program is covered by ST-002-SOP and ST-002-WIN-01.

3 ASSOCIATED DOCUMENTS

Document title	Reference (file name or path)
ORTA Registration/Randomization Program Approval Form	ST-002-AF-01
Implementation of Registration/Randomization	ST-002-WIN-01
External document on the minimization algorithm at the EORTC HQ.	J:\UNIT\Stat\3. Minimization\SOPs & descriptive documents

Forms are available from:

- Intranet, documentation section, Electronic Library of Quality Standard Documents page
- MS Word, File tab, New, My templates, forms tab, select the form

Work instructions (WINs) are available from Intranet, documentation section, Electronic Library of Quality Standard Documents page

4 DOCUMENT HISTORY

Version N°	Brief description of change	Author	Effective date
1.00	Initial release	Corneel Coens	25/02/2011
2	Consistency with other documentation and programs reviewed. Update of block randomization.	Corneel Coens	18 Aug 2015