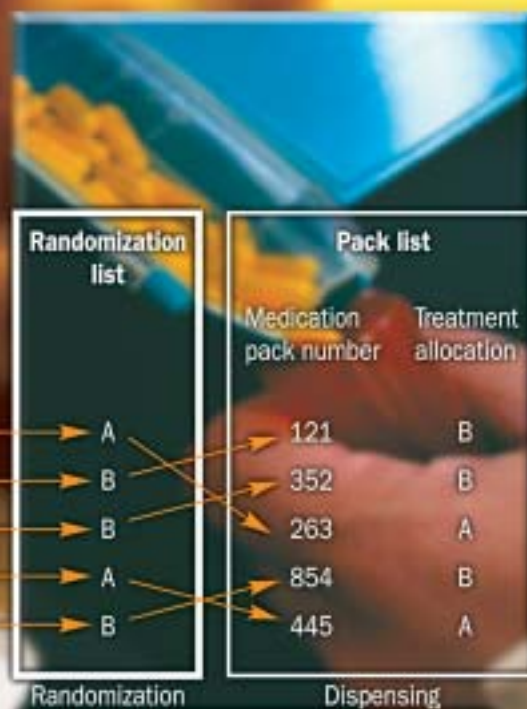


Forced Randomization

When Using Interactive Voice Response Systems

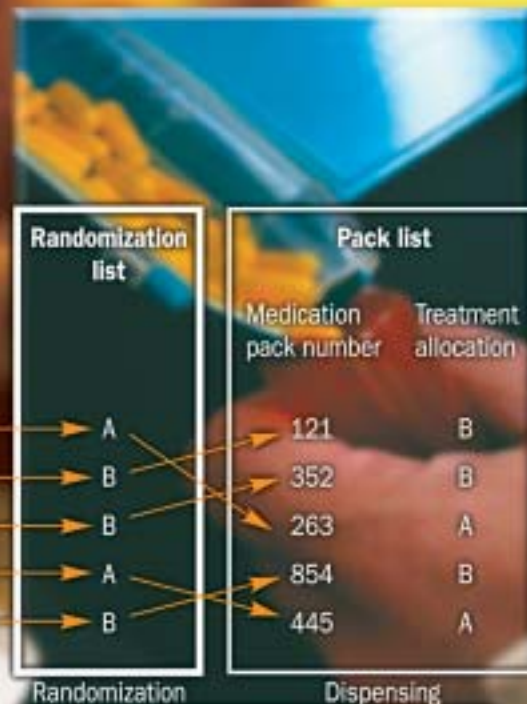


Damian McEntegart

Forced Randomization

When Using Interactive Voice Response Systems

Damian McEntegart



Forcing subject randomization may not be a by-the-book practice, but when combined with an IVR supply management system it may prevent unblinding and waste at the study site.

Interactive voice response systems are commonly used in clinical trials to manage the flow of trial medication supplies to sites and to manage the allocation of these supplies to individual subjects. Other advantages and uses include access to real-time information for trial managers, collection of diary card data directly from subjects, and as an aid to subject recruitment.¹

In this article we focus on the relationship between treatment randomization and supply management in IVR trials. As I will explain, in IVR trials substantial medication savings can be made using medication that is not labeled by subject number. This, however, does mean that it is possible for a site to run out of one particular type of medication but still have stocks of other treatments remaining. In this situation it is possible to restrict the randomization to the treatments for which stocks of medication are available. This has commonly become known as “forced randomization.” There are also other situations in which forcing can be used in trials managed by IVR.

Most statisticians find the concept of forcing anathema when they first encounter it. After hearing the issues discussed in this article, however, many do permit forced randomization to occur.

To appreciate the arguments surrounding forced randomization, it is necessary to have an understanding of the rationale for randomization and also how medication stocks are managed in a trial using IVR.

Randomization

Treatment allocation in comparative clinical trials is virtually always performed with some element of randomization. The two principles underpinning the design and execution of comparative clinical trials are minimization of bias and maximization of the precision of treatment effect estimates. Randomization, in conjunction with blinding (for blinded trials), is a key contributor to achieving precise and valid estimates of the treatment effect. Randomization prevents any conscious or unconscious selection bias by the investigator in the allocation of subjects to treatment groups. It provides a sound basis for statistical inference at the analysis stage of the trial. In contrast to any systematic allocation system, it tends to ensure balance between treatment groups on known and unknown factors that correlate with trial outcome independently of treatment.

With nonrandom allocation, there may be an underlying systematic reason why balance may not be achieved. To take an extreme example, a system that allocates according to the first letter of a subject’s surname may produce groups that are unbalanced on demographic factors because of the association of surname with country. For example, surnames beginning with “M” are particularly common in Scotland and Ireland due to names

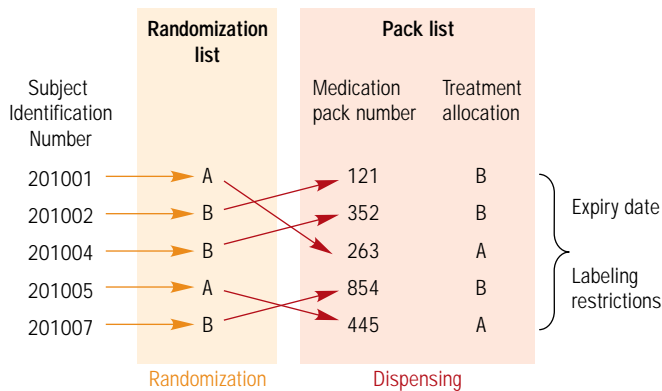


Figure 1. IVR randomization and medication dispensing.

beginning with “Mc” or “Mac.” Balance increases the precision of the treatment effect estimate and avoids so-called accidental bias.

The main method of randomization employed in clinical trials is allocation from a preprepared randomization list composed of random permuted blocks. For instance, in a trial of two treatments and a block size of B, half of the numbers within each block (B/2) are assigned to each of the treatments in a random order. Use of blocks helps to ensure that treatments are allocated in the planned ratios and helps comparability of treatment groups should subject characteristics vary over time. If there are stratifying factors (for example, gender and current smok-

Using IVR can reduce wastage that would traditionally occur when subjects withdraw from the study.

ing status) then the list is divided into sections, one for each stratum formed by the cross-classification of factor levels (in our example, male smoker, female smoker, male nonsmoker, and female nonsmoker). To prevent selection bias by the investigator, allocation should be sequential as stipulated by the ICH E9 Guidance which states that “The next subject to be randomized into a trial should always receive the treatment corresponding to the next free number in the appropriate randomization schedule (in the respective stratum, if randomization is stratified).”² The investigator can exercise selection bias if he knows or can guess the treatment allocation for the next subject. For instance, in an open-label trial stratified by center, treatment allocations towards the end of a block become ever more predictable. This knowledge may influence the investigator’s decision about subject entry into the trial.

If there are many stratifying factors, then the balancing properties of stratified randomization may break down because there are too many strata that have incomplete blocks at the end of the trial; this is termed “over-stratification.” In these situations, an alternative to the use of a randomization list is dynamic allocation in which the allocation of treatment to a subject is influenced by the current balance of allocated treatments and, in a stratified trial, by the stratum to which the subject belongs and the balance within that stratum. The most commonly used

dynamic allocation technique is minimization, which can cope with more stratifying factors for a given trial than randomization from a list. In minimization, allocation is made to the treatment that minimizes the imbalance calculated over the stratification factors. The ICH E9 Guidance states that an appropriate element of randomization should be retained. For example, the treatment that is optimal for balance is identified but then only allocated with a certain probability, for example a probability of 0.80, in a trial of two treatments. A recent comprehensive review of minimization is provided by Scott et al.³ Other dynamic allocation techniques, including some with slightly greater statistical efficiency than minimization, are detailed in a recent overview.⁴

Medication management in IVR trials

A complete description of this topic is provided by Dowlman.⁵⁻⁶

In IVR or Web-based trials, savings can be made by separating the randomization numbering system from the medication pack numbering system (as shown in Figure 1). At the time of randomization, the subject is assigned to a treatment corresponding to the next available number in the appropriate section of the randomization list. A pack of the treatment to be allocated is chosen from the stock at the site and the IVR system reads out the pack number to dispense. The choice of pack is best done randomly from the whole stock or the subset of stock with the earliest expiration date—the important point being that there should be no attempt to pick sequentially from an ordered list of pack numbers or else gaps in the list may partially unblind the investigator.⁷ If supplies are packed into units that cover durations shorter than the whole trial length, using IVR can reduce wastage that would traditionally occur when subjects withdraw from the study because units may be allocated to other subjects in the same treatment group.

Another supply saving from IVR comes from the use of automated inventory stock management that allows the targeting of stocks to the more productive recruiting sites. The higher recruiting sites cannot be identified with certainty before the trial begins, and so a system that adaptively focuses supplies to these more productive centers holds particular attraction. In trials managed by IVR, medication inventories at site are managed by algorithms that automate the generation of supply requests to the depot when stocks fall to an amount that does not meet the anticipated supply needs. Typically these inventory control algorithms are run every 24 hours. A “buffer stock” can be kept at site to meet the supply needs of new randomizations and replacements for damaged or lost stock. Once the site stocks fall to the minimum stock level (the “trigger level”) for a particular treatment, then this is recognized by the IVR system and a consignment request is sent to the depot to order the medication needed to restock all treatments to their maximum stock levels (“resupply level”). Restocking all treatments avoids issues of unblinding and reduces the number of consignments that need to be sent during the trial. The supply needs of subjects at post-randomization visits, as they progress through the trial, can be managed by an additional predictive algorithm that predicts when the subject will come back and ensures that there is sufficient stock at the site to meet the resupply needs of all ongoing subjects for a given period of time.

Trigger and resupply levels should be set according to the

needs of the particular trial. In the ideal world of unlimited budgets and medication, the trigger level should be set sufficiently high so that when the consignment order for a resupply is raised, there is still sufficient stock at site to cope with any subjects presenting for randomization while resupply from the depot occurs. If the depot is in the same country as the site, we typically allow between two and seven days depending on the country and depot. Site-specific trigger levels can be set according to the site's known or expected recruitment performance. But if medication is in short supply or particularly expensive, there is a motivation to set trigger levels and resupply amounts to the minimum level, so that there is less material wastage at nonproductive centers and for the trial as a whole. One strategy is to set trigger levels sufficient to deal with a certain percentage (for example, 99%) of situations determined by a Poisson recruitment process and the known distribution of delivery times. But

The choice of randomization method has ramifications for the setting of trigger levels and the probability of running out of stock of one particular treatment type.

however the levels are set, there is always a chance of running out of stock of one treatment type because of an atypical sudden surge in recruitment in a short period.

The randomization method choice has ramifications for the setting of trigger levels and the probability of running out of stock of one particular treatment type. If the randomization is not stratified by site, then there is no control over the balance of subjects randomized at a particular site in the period between an order being placed and delivered. For instance, in a three-treatment-group trial there is a one-in-nine chance that three successive subjects will be randomized to the same treatment group if there is no site stratification. In contrast, if the randomization is stratified at the site level, the probability is zero for a block size of three and 0.02 for a block size of six. Using a dynamic allocation scheme with site as a balancing factor is an alternative to randomization stratified at the site level. Here, the random element employed and the fact that the imbalance measure is typically calculated over a number of stratification factors means that runs of the same treatment allocation are also possible. Thus the trigger and resupply levels should be set taking into consideration the randomization method being used. The ICH E9 Guidance discusses the question of when stratification by site is appropriate.

Options when stock runs out

From the above it can be seen that it is possible that subjects will present for randomization without the site having supplies of all potential treatments. There are then three options that should have been predefined in the IVR system before the start of the trial:

Halt the randomization and instruct the investigator to contact the monitor.

Check what the next randomization should be, and if it corresponds to a treatment that is available at site, allow the randomization to proceed. Otherwise, halt the randomization as before. An aspect of this option is that if there is stratification involved, then the call has to proceed to a certain point to collect the stratification information. This may make the investigator more suspicious about the reason for stopping the call and consequently makes the possibility of partial unblinding more likely (see below). On the other hand, more subjects will be randomized as compared with the first option.

Allow the randomization to proceed and “force” the subject to be allocated to the next free number in the randomization list that corresponds to a treatment available at the site. Because of the nonsequential allocation of randomization numbers, the forcing option is counter to the ICH E9 Guidance and, for this reason, forcing will not be viewed favorably by some. Further discussion of five points is, however, warranted to provide the full picture.

- The main focus of the ICH E9 Guidance is to prevent selection bias, and this is not the case with forcing.
- If the subject cannot be randomized, the investigator and subject may be greatly inconvenienced. Further, if the randomization visit cannot be rescheduled, the patient may not enter the trial at all—this is especially irritating if the patient has completed a screening/run-in period. If the investigator suspects the true reason, then some may pick a pack for allocation and inform the IVR vendor after the event. Some investigators have argued that once a subject has been screened, it is unethical not to proceed with treatment, particularly if any trial procedures have been undertaken.
- Allowing forcing prevents partial unblinding where, if the randomization does not proceed, the investigator may make certain deductions about the packs currently in the site's inventory; that is, not all treatment types are represented. In a trial of many treatments, this may not be too great an issue. In a two-treatment-group trial, the problem is more acute in that the investigator may deduce that the packs “on the shelf” are all of the same treatment. Should the identity of one of these packs subsequently become known to the investigator, through an emergency code break or through guesswork based on observed pharmacological effects, then all the stocks currently in the inventory are unblinded. If some of these packs are used for the resupply of existing randomized subjects, then further deductions may be possible. If forcing is not permitted, in the ideal world the potentially unblinded stocks should all be withdrawn and replaced as soon as the randomization call is blocked; this happens very rarely.
- The extent to which partial unblinding is a real or theoretical problem depends on the situation. Generally the investigator would have to note or remember the numbers or shelf position of the original medication for there to be an issue. The latter is more likely if the investigator stores the medication sorted into pack number order. When investigators contact their monitor, the monitor often refers the investigator to the IVR vendor help desk. Our experience is that it is difficult to satisfy investigators as to why they cannot randomize subjects when they know they have supplies available and, although they cannot be told the nature of the problem, they often will deduce it. Whether they then use this information is unknown.

- If many randomizations are forced, then we may lose the protection of randomization in ensuring balance on prognostic factors if there is any pattern to the sites that run out of medication. But if there are relatively few forced randomizations, and further if the treatment that runs out is quickly resupplied for potential allocation to future subjects, then there should be no problem.

If the randomization is forced, then there is a subsequent decision to be made about whether to allocate the next subject the “skipped” randomization number or to leave this number unallocated (in fact, due to the blocking several randomization numbers may have been skipped but the same principle applies). If balance is important then it is best to “backfill” the number, and this is what is most commonly done. Alternatively, if the skipped numbers are left unused, then the subsequent audit trail is easier in that a gap will indicate forcing.

Some statisticians seem more comfortable using forcing with a dynamic allocation technique such as minimization. When not all treatments are available at site, the algorithm is run as usual but only the available treatments are passed through. In my opinion, the principle is no different to forcing from a random-

Using a dynamic allocation technique with forcing is no different than forcing from a randomization list, and comfort derived from the fact that the letter of the ICH E9 Guidance is not strictly violated is spurious.

ization list as not all treatments have a chance of being selected. Any comfort derived from the fact that the strict letter of the ICH E9 Guidance is not strictly violated is spurious.

Forcing as part of trial design

The above discussion relates to forcing due to unplanned lack of medication at the center. There are three occasions where forcing in IVR is more intentionally incorporated into the design and these are outlined below.

Forcing at site using a double randomization. We have used this strategy in the case of many treatments, scarce supplies, and relatively low recruitment at each site—typically a certain proportion of sites would only recruit one or two subjects or even fail to recruit any subjects. In this situation sending a full set of treatment supplies represents a substantial amount of wastage.

To reduce wastage, two separate randomization lists are employed. A first randomization list is prepared using the smallest block size, which will be the sum of the allocation ratios. For example, a block size of seven will be used for a trial of seven treatments with an equal allocation ratio; a block size of eight would be used for a trial of five treatments to be allocated using a 3:2:1:1:1 ratio. As sites are activated in IVR, they are sent supplies corresponding to a fraction of the block size. For instance, in the trial of seven treatments, the first site activated may be sent treatments corresponding to the first three entries on the

randomization list. The second site activated would be sent the next three entries on the randomization list and so on. This means that a site could be sent two packs of the same treatment drawn from separate blocks. Randomization is performed from an entirely separate randomization list and uses forcing to allocate the subject the next available randomization number corresponding to a treatment that is available at the site in question. Randomization numbers corresponding to non-available treatments are skipped but are available for subsequent subjects for backfilling. As subjects are recruited at a site, resupplies are initiated using the first randomization list. By forcing allocation from a balanced list, this technique will result in obtaining the best overall study treatment group balance in the face of limited supplies and a minimal wastage of medication.

Again, this scheme is counter to the ICH E9 Guidance. Arguably though, the protection against predictability and selection bias is actually increased by the scheme as, even if the investigator knows the block size and past treatment allocations, this does not help him or her predict the medication for the next subject to be randomized. In permuted blocks randomization allocations towards the end of the block become ever more predictable if the investigator knows or can guess the previous allocations. The scheme is easily extended to allow for stratification factors.

Forcing for balance. Sometimes we force to try and maintain balance at a particular factor, usually center. The idea originated with Zelen.⁸ Treatments are assigned from an unstratified, blocked randomization list as usual, except that if the assignment would result in imbalance above a certain magnitude, D , within a center, then the next allocation is made to the treatment group with the lowest allocation. The value of D can be varied randomly for each subject or can be fixed in advance provided it is not revealed to the sites. The scheme can be extended to cover randomization stratified for prognostic factors other than site. A random element can be applied so that forcing for balance only occurs with a high probability, for example 0.80. The scheme can be used with more than two treatments basing the criterion D on the range of treatment allocations.

This technique can be useful to avoid extreme imbalances within a site where it is not desirable to stratify by center due to concerns about overstratification, for example low numbers at each site relative to the number of treatment groups and/or other stratification factors. Provided that the entry of subjects into the trial is random, the net effect is that treatment allocation is random within centers.

Accounting for screening drop-outs. To limit medication wastage in studies with a screening or run-in period, a variant of the prediction method can be used. As the subject is enrolled into the run-in period, the IVR system allocates the next entry from the randomization list and raises an order to send the appropriate medication to the site for use should the subject be successfully randomized. If the randomization is stratified by site then there is no need for forcing with this scheme. Alternatively, if the scheme is not stratified by site and allocation is from a central, unstratified randomization list, then the subject is effectively being randomized at screening. This is in violation of the ICH E9 Guidance which states that “The appropriate number and associated treatment for the next subject should only be allocated

when entry of that subject to the randomized part of the trial has been confirmed." If the screening period is of variable length, then even if all subjects entering the run-in successfully enroll into the randomized phase of the trial, they will not necessarily receive the same treatment they would have been randomized to upon confirmation of their entry to the trial. In the more usual case where some subjects entering the run-in period are not randomized, then a form of forcing is that the IVR system reserves

Provided there are not too many instances and that, if necessary, the primary results can be confirmed by sensitivity analyses, trials that use forcing can be accepted as confirmatory trials.

the treatment supplied for the run-in failure for the next subject to enter the run-in period at that site; the first available randomization number that corresponds to the treatment is assigned. There are no unblinding issues with this strategy. The method works best when the proportion of run-in drop-outs is low.

Other aspects of forcing

An example. Hamilton describes a trial of four treatments to be allocated to 500 subjects at 50 hospitals.⁹ Medication was in short supply. Treatment allocation was to be balanced over three prognostic factors and within hospital. The allocation scheme used was minimization. The inventory scheme used was an initial shipment of one kit for each treatment and a resupply with two new kits when two kits remained. The minimization algorithm allocated to treatment choosing from those available at site at the point of allocation. Simulations showed that the optimal treatment identified for balance by the minimization algorithm would be available for assignment in 91% of allocations. In the remaining 9% of assignments, the second-best treatment for minimizing imbalance was always available at site. The simulations showed that of 14 other resupply schemes investigated, only one scheme guaranteed that the first-choice preference of the minimization algorithm was always available for assignment. This scheme required a mean of 1095 kits to be shipped compared to a mean for the chosen scheme of 630. Thus the restricted assignment employed in this study potentially reduced the amount of kit wasted from 119% of the amount used to 26%.

Regulatory perspective. Clients or colleagues have asked a number of European and FDA statistical regulators about this issue. The consensus seems to be that trials using forcing can be accepted in confirmatory trials provided there are not too many instances and that, if necessary, the primary results can be confirmed by sensitivity analyses. In this respect the situation is the same as that which applies for protocol violators. If forcing is permitted, the statistical analysis plan should define which analysis populations the forced subjects are to be included in.

Discussion

I have identified different types of forced randomization and illustrated the consequent medication savings that can be achieved. Assuming that subject entry to the trial is random and the proportion of forced randomizations is not too high, then all schemes retain the properties of unforced randomization relating to avoidance of selection bias and the tendency to produce treatment groups that are balanced for prognostic factors. Indeed, by preventing partial unblinding, forcing may reduce selection bias compared to the situation where forcing is not allowed.

In certain circumstances it could be argued that forced randomization can be ethical in the sense of collective ethics that consider the advancement of public health. If medication is in extreme short supply for unavoidable reasons, then the choice may be between a trial of longer duration conducted at a few centers and a trial that employs forcing at a larger number of centers. Important medical advances may be made earlier from the trial of shorter duration.

But whatever the arguments, forcing is counter to the strict letter of the ICH E9 Guidance. Individual trialists should therefore consider the arguments carefully before making the decision as to whether forcing is beneficial for their particular trial.

References

1. Bill Byrom, "Using IVRS in Clinical Trial Management," *Applied Clinical Trials*, October 2002, 36–42.
2. International Conference on Harmonisation, E-9 Document, "Guidance on Statistical Principles for Clinical Trials," *Federal Register* 63 (179) 49583–49598 (1998).
3. N.W. Scott et al., "The Method of Minimization for Allocation to Clinical Trials: A Review," *Controlled Clinical Trials*, 23 (6) 662–674 (2002).
4. D.J. McEntegart, "The Pursuit of Balance Using Stratified and Dynamic Randomization Techniques: An Overview," *Drug Information Journal*, 37 (3) 293–308 (2003).
5. N. Dowlman, "Intelligent Medication Management—Using IVR to Optimise the Drug Supply Process," *Pharmaceutical Manufacturing and Packing Sourcer*, (Summer) 24–28 (2001).
6. N. Dowlman, "The Cost Benefits of Using IVR Systems in the Supply Chain," *Pharmaceutical Manufacturing and Packing Sourcer*, (Autumn) 74–79 (2001).
7. S. Day, Letter to the Editor, *Controlled Clinical Trials*, 22 (3) 264–265 (2001).
8. M. Zelen, "The Randomization and Stratification of Subjects to Clinical Trials," *Journal of Chronic Disease*, 27, 365–375 (1974).
9. S.A. Hamilton, "Dynamically Allocating Treatment When the Cost of Goods is High and Drug Supply Is Limited," *Controlled Clinical Trials*, 21 (1) 44–53 (2000).

Damian McEntegart, MSc, is statistics and medication forecasting manager for ClinPhone Group Ltd., Lady Bay House, Meadow Grove, Nottingham, UK, +44 115 955 7333, fax +44 115 955 7555, email:dmcenteg@clinphone.com.