

## ALLOCATION OF PATIENTS TO TREATMENTS

From the Therapy chapter for the 3<sup>rd</sup> edition of Clinical Epidemiology, by DL Sackett  
17 April 2004 (day 108)

### Allocation Check List:

1	<input type="radio"/>	Double-check your system for registering and keeping track of every eligible consenting patient as soon as you identify them.
2	<input type="radio"/>	Set up a system for obtaining prognostic information about them if necessary.
3	<input type="radio"/>	Select an allocation method (stratified randomized blocks, minimization, factorial, etc).
4	<input type="radio"/>	Select some person or team to implement the allocation.
5	<input type="radio"/>	Set up safeguards to establish allocation concealment at the start of the trial.
6	<input type="radio"/>	If a blind RCT, set up systems to maintain both allocation concealment and blinding throughout the trial.

After a participating neurologist in the RRPCE trial completed an eligible and consenting patient's admission forms, he telephoned our central office, identified himself and the patient, and told us the vascular territory for the patient's TIAs (carotid, vertebro-basilar, or both) and whether they had any neurological residua. Our central office staff then consulted a previously generated (by computer) randomization schedule for that neurologist and those clinical features (balanced every 4 patients within each stratum), and gave the neurologist a randomly generated four-digit number that identified pre-packaged supplies of the two study drugs or their corresponding placebos already on hand in his hospital's pharmacy. We asked patients to avoid non-trial aspirin, and made arrangements with the clinical laboratories at each center to withhold telltale laboratory results (because sulfinpyrazone lowers serum uric acid).

Note that this set of tactics achieved four important objectives:

1. The neurologist could not influence the patient's assignment to a specific treatment, because we had concealed our allocation scheme.
2. We could guarantee a good prognostic factor balance between the treatment groups by stratification before randomization.
3. We set the stage for blinding both the neurologists and their patients as to who was receiving which of the four treatment regimens.
4. We established an audit trail for every randomized patient.

Because of their interconnectedness<sup>1</sup>, we won't try to consider these items one-by-one. Instead, we will integrate them as we answer some important questions about allocation in RCTs. The objective in all this is to help you collaborate with your statistician-co-investigator in making decisions about how to allocate your trial patients to their treatments.

### Why it is vital to conceal the assignment of patients to treatments

Systematic reviews of RCTs have shown that when clinicians know ahead of time which treatment their next eligible and consenting patient will receive, they may (consciously or unconsciously) enter patients with lower risk and/or higher responsiveness into the experimental treatment group. As a result, the trial can become biased in favor of experimental therapy from the start. For example, Ken Schulz has led empirical studies showing that trials that failed to conceal their assignment schemes tended to exaggerate treatment effects (although with scope for bias in either direction)<sup>2</sup>. However, this assignment of low risk/highly responsive patients to experimental treatment is not universal. Regina Kunz, Gunn Vist, and Andrew Oxman performed a systematic review that uncovered examples of underestimating as well as overestimating treatment effects in RCTs with inadequate concealment<sup>3</sup>.

### How to conceal the assignment of the next patient

The mechanics of concealing patient allocation to treatment vary widely, but all of them rely on assignment by an external source that neither clinicians nor patients can influence or even know about (alas, locally held allocation lists can be leaked or burgled, and “sealed” envelopes can be opened prematurely or held up to a strong light). The allocation strategy could be as simple as a coin-toss by a third party (but remember that 100 coin tosses usually contain 6 heads or tails in a row at some point in the sequence!). Nowadays allocation is often performed by a central randomization service (such as a 24/7 telephone number) that employs a computer-based randomization algorithm, with or without provisions for the prognostic stratification and minimization approaches discussed in the next section. Many statistics texts contain random number tables, and statistical software often includes random number generators. Your choice of the specific strategy to use in your trial should arise out of discussion with your principal statistician.

In drug trials it is usual for the allocation system to safeguard concealment by assigning patients to specific previously-assembled and identical pill containers kept in stock at the study sites. You should label each of these containers in a fashion that defies identification of their contents. For example, we “numbered” each container in the RRPCE trial with four random digits.

Allocating roughly equal numbers of patients to experimental and control therapy, both overall and in sub-groups, boosts the power of the RCT (confidence intervals around, say, absolute risk reductions are smallest when treatment groups are identical in size)<sup>1</sup>. Roughly equally-sized groups also increase the study’s credibility before clinical audiences. Accordingly, randomization schemes typically employ “balancing” or “blocking” exceptions to strict random allocation. For example, suppose that you want to balance the numbers of experimental and control patients at the end of every 4 who are entered (a “block” of 4). Then, if both the first and second patients in your block of 4 were randomly assigned to control therapy, you would not randomize the third and fourth patients, but would assign both of them to experimental therapy. Although blocking is tidy, it can be risky. If you were “blocking” within admitting teams or sites and they got wind of your blocking scheme and kept track of patient assignments, concealment would be lost and they could determine the next patient’s treatment. For this reason, you should keep your blocking plan secret from study clinicians, and some investigators randomly switch between blocks of different numbers of patients.

When clinicians or patients are only slightly uncertain that a new treatment does more good than harm, some authors suggest changing the allocation ratio of experimental:control assignments from 1:1 to 2:1 or even 3:1. If you decide to do this, you must remember that, to obtain the same confidence interval around the treatment effect, a 2:1 trial requires 12% more patients overall and a 3:1 trial requires 33% more patients (so that a 3:1 trial doesn’t reduce the number of control patients by 50%, but only by 33%)<sup>4</sup>.

### **Why prognostic factor balance is important**

Just as Figure 1 of any RCT report typically provides a “flow-chart” of patients as they progress through the trial, Table 1 of any RCT report typically provides a “baseline” comparison of experimental and control patients upon admission to the trial. The key entries in this table are factors known to affect study patients’ risks of the outcomes of interest or their likely

---

<sup>1</sup> But not too identical. Ken Schulz and colleagues examined 206 obstetrical and gynaecological RCTs and found that differences in the numbers of treatment and control patients were far smaller than would be expected by the play of chance, raising suspicions that some investigators were interfering with randomization. [Schulz KF, Chalmers I, Grimes DA, Altman DG. Assessing the quality of randomization from reports of controlled trials published in obstetrics and gynecology journals. JAMA 1994;13:125-8.]

responsiveness to the experimental treatment. You will enhance readers' confidence in your trial result when you achieve a close balance in these prognostic factors.

Many RCT reports make the mistake of testing the statistical significance of any difference in each prognostic factor between experimental and control patients. In small trials, biologically important differences will often be statistically non-significant; in large trials, biologically trivial differences will often be statistically significant. As the following example will show, the issue is more one of credibility than validity.

### Threats to trial credibility from prognostic factor imbalances

By its very nature, randomization must inevitably occasionally lead to random but big differences between treatment groups in randomized trials. The results on the credibility of the affected trials can be devastating. For example, an excellent team of trialists designed and executed an RCT<sup>5</sup> that, among other objectives, asked the question: "Among patients with type II diabetes, does an oral hypoglycemic agent (tolbutamide), when compared with placebo, reduce the occurrence of nonfatal vascular complications and death?" They employed simple random allocation (with no prior stratification), and a comparison of vascular risk factors between the tolbutamide and control groups at the start of the trial looked like Table 3-6-1.

Table 3-6-1: Risk factors for vascular complications and death at the start of the trial

Risk Factor	Placebo Group	Tolbutamide Group
Definite hypertension	30%	37%
History of digitalis use	4.5%	7.6%
History of angina pectoris	5.0%	7.0%
Significant ECG abnormality	3.0%	4.0%
Elevated cholesterol	8.6%	15%

That all 5 of these risk factors were more common in the tolbutamide group is unusual in an individual trial (sign-test  $P=0.0625$ ), but inevitable in a large group of trials. Nonetheless, this non-statistically significant imbalance became a key target for the drug's manufacturer (and its hired consultants<sup>6</sup>) when it was found that the active drug produced a 44% increase in total mortality on the raw data. In truth, however, this baseline imbalance didn't affect the study's conclusions. Multivariate modeling with logistic regression (to control for these imbalances) reduced this increase by just 2%, to 42%. Moreover, an independent team of the best statisticians in the land "adjusted" for important baseline differences and upheld the study conclusion<sup>7</sup>. For some readers, however, the damage had already been done, and the credibility of this important RCT suffered.

### Achieving prognostic factor balance by stratification prior to randomization

Our research groups usually employ one of two "mixed" allocation strategies to prevent this distracting situation from happening. The first of them employs prognostic stratification before randomization. It begins by stratifying each patient for the presence/absence of important prognostic factors, or for whether they lie above or below a guesstimate of the median value for a continuous prognostic factor (such as blood pressure) or effect modifier (such as age). Or, if we are performing a simultaneous explanatory-pragmatic trial, we stratify by whatever pre-randomization evaluation identified whether a patient should go in the explanatory or management arm. Then, allocation occurs separately within each stratum.

For example, in the RRPCE trial we stratified for three presumed sites of ischemia (carotid, vertebrobasilar, or both) and for the presence or absence of residual signs and symptoms

following the qualifying TIA. We thus had 3 x 2 or 6 strata from which we randomized patients to their study regimens. We employed a separate randomization schedule, balanced every four patients, for each stratum (with an independent randomization schedule for each center). The resulting baseline balance in these prognostic factors for men in our trial appears in Table 3-6-2

Table 3-6-2: Baseline balance for prognostic factors among men in the RRPCE study

Site and Residuuum	Sulfinpyrazone	Aspirin	Both	Neither
Carotid, with residua	27%	29%	28%	30%
Carotid, no residua	35%	37%	41%	34%
V-B, with residua	10%	6%	8%	7%
V-B, no residua	15%	19%	17%	13%
Both, with residua	9%	3%	4%	2%
Both, no residua	4%	5%	3%	13%
Totals	100%	100%	100%	100%

If you run your eye along the rows, you'll see good balance for these 2 prognostic factors. Every treatment group had the highest % for one factor, and no group had the highest rate for more than 2 of them. No wonder, then, that baseline balance was never an issue in the interpretation of this trial, even by its detractors.

### Achieving prognostic factor balance by minimization

The second "mixed allocation" strategy we have used to achieve prognostic factor balance is named "minimization"<sup>8</sup>. It is particularly useful for achieving good balance in small trials. You begin by identifying the prognostic (or simply cosmetic) factors you want balanced. Then you dichotomize them, either at some estimated mid-point or at some clinically sensible break-point.

For example, in one of our compliance trials we had just 38 non-compliant, uncontrolled hypertensive steel workers in whom we wanted to test a set of behavioural strategies for their effects on compliance and blood pressure control<sup>9</sup>. We were concerned about balance (for reasons of credibility as well as confounding), so we began by identifying three factors we thought needed balancing:

1. The level of their diastolic blood pressure (upper vs. lower halves of their overall distribution of diastolic blood pressures)
2. Their past compliance 6 months into their treatment for hypertension (upper vs. lower halves of their distribution of % compliance by pill-count).
3. Whether they had previously gone through a special education programme for hypertensives

We assigned arbitrary scores of 1 each to the upper half of diastolic blood pressure, the lower half of compliance, and having received education. Thus, a study patient with "upper half" blood pressure, "lower-half" compliance, and who had not received education would score 1+1+0=2 points.

We then allocated patients according to a set of rules that minimized the differences in total points between the experimental and control patients. Although we've long since thrown away the actual allocation schedule for this 1974 compliance trial, the schedule shown in Table 6-3 is an accurate representation of how we handled each patient:

Table 3-6-3: Allocation by minimization

Patient #	Score	Allocated by:	Running Score Totals	
			Experimental	Control
At the start			0	0
1	2	Randomization	0	Randomized here, so running score for control patients = 2
2	1	Minimization	Minimized here, so running score for experimental patients = 1	2
3	2	Minimization	Minimized here again, so 1+2 = 3	2
4	1	Minimization	3	Minimized here, so 2+1=3
5	0	Randomization	Randomized here, but 3+0 still = only 3	3
6	3	Randomization	3	Randomized here, so 3+3=6
7	2	Minimization	Minimized here, so 3+2=5	6
etc	etc	etc	etc	etc

- Patient #1 was randomized (she had a score of 2, and was randomized to the control group).
- Patient #2 had a score of 1, and to minimize the difference in total scores between the two groups, he was allocated to the experimental group.
- Patient #3 had a score of 2, and minimization allocated him to the experimental group as well.

And so on, as shown in Table 3-6-3. Note that every time the running scores are tied, the next patient was randomized.

By employing minimization, we achieved excellent balance for these prognostic factors in our small compliance trial (its execution and positive result appears as the scenario for the section on small trials on page xx). Moreover, in part because of the correlation between the factors we chose and age, height, weight, and symptoms, we also achieved excellent balance for these latter features.

There are a few other “chance” methods for assigning patients to treatments, but we won’t discuss them here; they are nicely reviewed in the Lancet series from Ken Schulz and David Grimes<sup>10</sup>.

### Should minimization replace randomization in all trials?

Although it may sound heretical, a strong argument can be made for choosing minimization over randomization in all RCTs, even the very large ones. This argument rests on a series of realizations. First, randomization serves two purposes: concealment and balance. Second, there

are other tactics (such as assignment through a central facility) for achieving concealment. Third, as we've just shown you, minimization is a better method for achieving balance for known prognostic factors.

However, that leaves the unknown prognostic factors. They are ignored in minimization, but on average balanced by randomization. Because of this potential imbalance in important but unknown prognostic factors, many methodologists are uncomfortable abandoning the "gold standard" of random allocation for the "platinum standard" of minimization<sup>11</sup>.

### Allocation of patients to multiple treatments: the Factorial Design

Thus far, we have described allocation to a single experimental group and a single control group. But, what if you want to study two drugs rather than just one, as in the RRPCE trial? Although you could perform a 3-arm trial (Drug A, Drug B, and Placebo) there is another design which gives the same information with fewer patients, and as a bonus may even detect interactions between the two drugs.

This is the factorial design in which patients (following stratification) are randomized (or minimized) to one of four groups as shown in Table 3-6-4.

Table 3-6-4. A factorial design

		Treatment A		Efficacy of S
		Active	Placebo	
Treatment S	Active	a	b	a+b vs. c+d
	Placebo	c	d	
Efficacy of A		a+c	vs. b+d	

As long as there is no "interaction" between the 2 drugs (such that one of them works better or worse in combination with the other than it does when given by itself), the efficacy of both treatments are determined "at the margins." That is, the efficacy of Treatment S is accurately determined by comparing outcomes for the combined arms (a+b) vs. (c+d), and the efficacy of Treatment A is accurately determined by comparing outcomes for the combined arms (a+c) vs. (b+d). As a result, it becomes possible to test 2 drugs for the price (in terms of sample size) of 1, which is why the factorial trial is often recommended by methodologists<sup>12</sup> and becoming increasingly popular among trialists. As noted at the start of this paragraph, however, when the drugs interact (such that their effects are not additive) it becomes inappropriate to assess their efficacy at the margins, and the sample size advantage is lost (We'll come back to this issue in the section on Analysis).

## Why blinding is important, and how to initiate it.

In an ideal RCT, the experimental treatment is given to everyone in the experimental group and to no one in the control group. Moreover, both groups are treated equally in all other respects. Finally, their outcomes are ascertained and reported with equal vigour and accuracy. These ideals are at risk whenever the treating clinician or patient “breaks the code” and learns who is receiving which treatment. The risks of “unblinding” are three:

1. Contamination of the control group: When either the clinician or the patient is pretty certain that the experimental therapy is superior, one or both of them may take steps to ensure that control patients receive it. The result of this contamination is a decrease in any difference in outcomes between the two groups. The damage done by contamination adds a methodological justification to the ethical justification for employing the “uncertainty principle” in recruiting patients for RCTs. As discussed in the “Principles” discussion in this section of the chapter, unless both clinician and patient are genuinely uncertain which treatment is better, the patient should not be enrolled in the trial.

2. Unequal co-intervention: New treatments are not tested in isolation, and patients in RCTs can (and, in the case of previously validated treatments, must) receive a wide array of other (we’ll call them “ancillary”) treatments that will, on average, favourably affect their outcomes. If these co-interventions are unequally applied to, or complied with by, experimental and control patients, it may become impossible to decide whether any end-of-study differences in their outcomes are due to the experimental treatment or to unequal co-intervention. The problem of co-intervention is especially troublesome when it is also an outcome used to assess the efficacy of experimental therapy (such as the decision to hospitalize or operate on a study patient).

3. Unequal ascertainment of outcomes: Some trial outcomes (such as total mortality) are unambiguously “hard.” That is, they are easy to ascertain, and even unblinded observers of a given study patient are unlikely to disagree about whether they have occurred. But what about “softer” but important outcomes (such as discreet clinical events that can occur in mild form, symptoms, function, or quality of life)? When blinding is lost, the ascertainment of these outcomes may be affected by study clinicians’ or patients’ preconceived notions about efficacy. When study patients or their clinicians know which treatment they are receiving, they may be followed more or less closely for “soft” outcomes, and their symptoms played up or ignored. Moreover, the threshold for carrying out definitive diagnostic testing may differ. Finally, a patient’s knowledge of their treatment (a “new and promising drug” or “just a sugar pill”) can affect not only their symptoms, but also their quality of life.

As shown above, concealed allocation and blinding, if protected throughout the trial, permit you to avoid contamination, co-intervention and unequal ascertainment of outcomes. We’ll describe some tactics for maintaining both in later stages of a trial as they arise.

You may have noticed that we never employ the term “double-blind” in this chapter. That’s because a team led by PJ Devereaux documented that this term has several meanings for both trialists and clinicians.<sup>13</sup> As a consequence, it cannot accurately describe the “blind” status of the (at least) 7 sorts of individuals<sup>2</sup> who are involved in any clinical trial.

We also haven’t replaced the term “blinding” with the euphemism “masking.” Readers who wonder why or prefer the latter term should ponder whether they’ve ever seen a party mask that didn’t prevent blindness by providing holes through which the wearer could see what was really going on.

---

<sup>2</sup> Trial patients, their clinicians, data collectors, outcome assessors/adjudicators, data analysts, trial monitoring committee members, and manuscript writers.



## **Should you test patients' and clinicians' blindness during and after your trial?**

Some writers about RCTs suggest testing clinicians and patients for “blindness” during and after trials. We don't, for 2 reasons. First, during the trial, we want our clinicians and patients to focus on following the protocol and taking study medications as prescribed. We don't want them distracted by “games” or detective work that might break the code<sup>14</sup>. Second, asking patients or their clinicians after a trial to guess which treatment each of them received is not a test for blindness. In fact, it confounds bad blinding with good hunches about efficacy. If that last sentence is either totally mystifying or terribly tantalizing, skip to page xx and learn more about it before proceeding further with this one. In summary, we work very hard before starting a trial to establish and test blindness in its “pilot” phase, but we never test for blindness during or after our trials.

## **What should you do when blinding is impossible?**

Finally, blinding is impossible for lots of trials. “Mock” surgery, such as skin incisions alone vs. full arthroscopic knee surgery<sup>15</sup>, is rarely employed, and its ethics hotly debated when it occurs<sup>16</sup>. Placebo physical or psychological therapy is difficult to design and apply. As a result, tactics for preventing or minimizing the effects of the absence of blinding in RCTs are very important. They include rigorous, all encompassing ancillary treatment protocols (to prevent unequal co-intervention), equally intense follow-up of experimental and control patients, and the employment of blind, external outcome adjudicators (to reduce biased outcome assessment). These tactics will be discussed in subsequent sections of this chapter.

## **Why an audit trail is important, and how to create one**

For scientific validity, and even more so for clinical credibility, you must be able to account for every patient who was allocated to treatment in your trial. This accounting has to include patients who refused their treatment allocation or left the trial for any reason, because they are essential members of any intention-to-treat analysis. Registration at allocation also permits you to start an individual patient file. Not only can you add all pertinent follow-up information to this file. You can also employ it to schedule their follow-up visits and to remind their clinicians if they fail to show up.

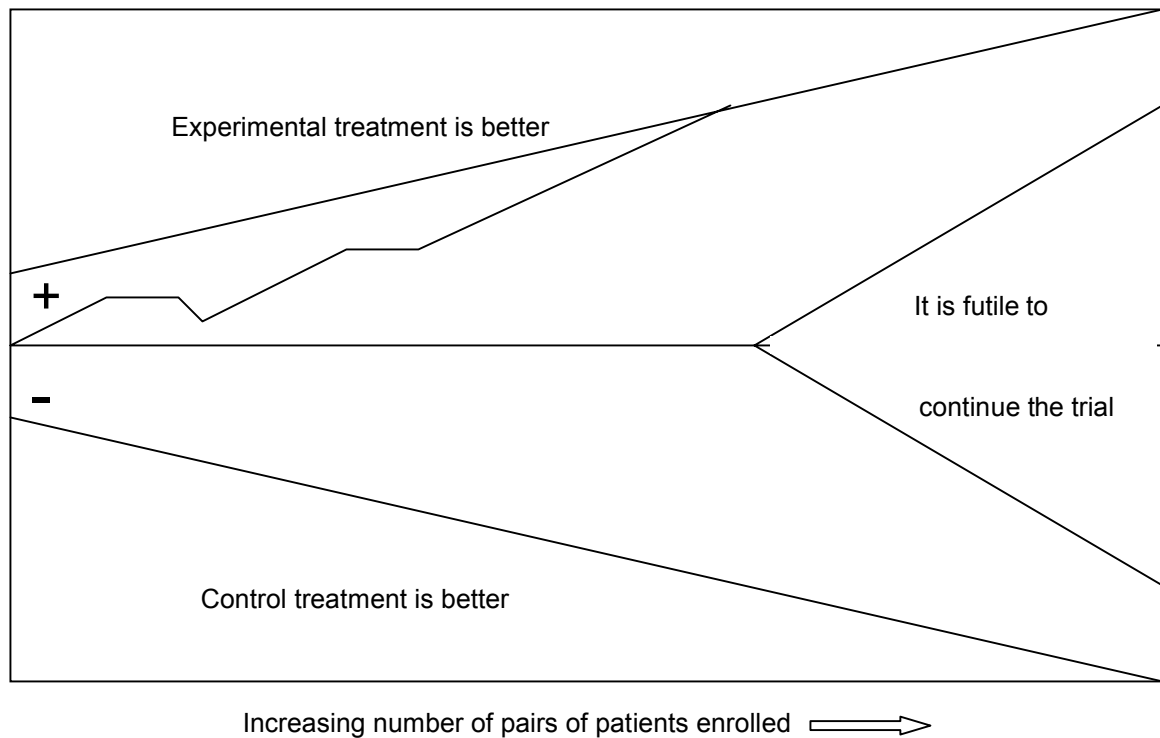
## **Alternative allocation strategies**

There are four other allocation strategies that are employed by some trialists, some times. We've never used them, so we won't pretend to know their ins and outs. Instead, we will refer you to other sources that describe them in greater detail.

1. **Paired sequential trials:** When outcomes occur quickly and eligible patients present in a steady stream, some trialists randomize the first member of a pair (sometimes after matching them for important prognostic factors), and assign the second member to the alternative therapy. The outcomes are assessed within each pair, assigning a “+” to the pair if the experimental patient fared better, a “0” if they fared the same, and a “-” if the control patient fared better. The results for successive pairs are then put into a graph such as shown in Figure 6-1, with enrolment of successive pairs along the X-axis their outcomes along the Y-axis. A “+” result raises the cumulative results line one unit along the Y-axis, a “-” result lowers it one unit, and a “0” result keeps it on the level. Statistical boundaries can be constructed such that, when they are crossed, a conclusion can be drawn that experimental therapy is efficacious or harmful, or that it is futile to continue the trial. In the example depicted in Figure 3-6-1, the experimental therapy ultimately

crosses the statistical boundary for efficacy. You can read more about sequential designs in Curt Meinert's book<sup>17</sup>.

Figure 3-6-1: A sequential trial



2. Adaptive allocation strategies based on outcomes: Some trialists have generated methods for altering the allocation ratio based on the outcomes of previously admitted patients. Thus, if the last patient allocated to experimental therapy did well, the probability that the next eligible patient would be allocated to experimental therapy would rise. These “play-the-winner” strategies are appealing to some trialists, but are at risk of revealing the likely allocation of the next patient. In addition, they are not well-suited for long-term trials with substantial delays between entry and outcome. Again, we refer you to one of the RCT texts for more about adaptive allocation strategies<sup>18</sup>.

3. Cross-over trials: When the outcomes of interest in a trial are symptoms, functional capacity, or other outcomes that produce no permanent changes in study patients, you can reduce “noise<sup>19</sup>” and your sample size requirement by giving both the experimental and control treatments to every patient, simply randomizing the order in which they are applied. Although ideal for some situations (see Gordon Guyatt’s description of “N-of-1” trials later in this chapter), you can’t determine whether the treatment given first has any permanent (“carry-over”) effects until the trial is over. If it does display this “order effect,” the power of the cross-over design is reduced to that of a two-group RCT. Moreover, in infertility trials where pregnancy is the outcome, a McMaster team showed that cross-over trials overestimated odds ratios for efficacy by 74%<sup>20</sup>. Stuart Pocock devotes a chapter to cross-over trials in his RCT text<sup>21</sup>.

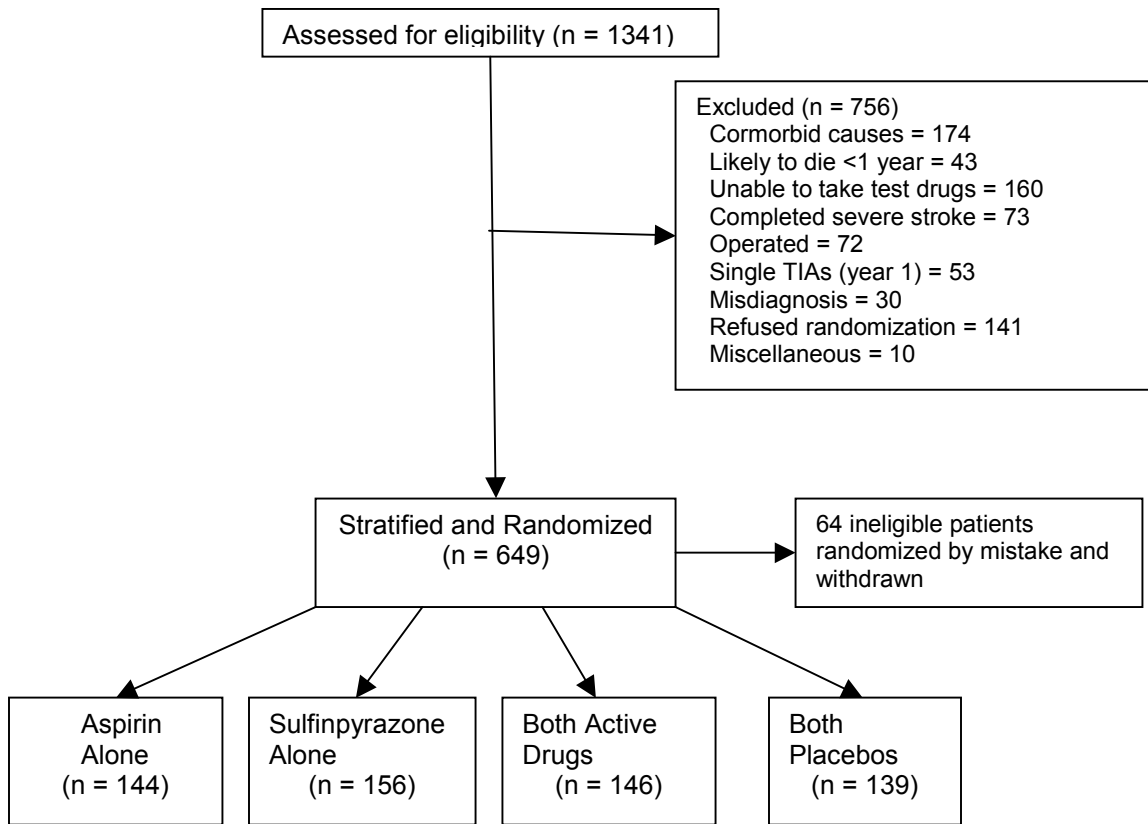
4. Randomized consent (“Zelen”) trials: Marvin Zelen introduced an allocation strategy in which eligible patients are randomized to one of two groups before seeking their consent<sup>22</sup>. Group 1 are offered the experimental treatment and undergo full RCT informed consent. Group 2 are offered current standard therapy and undergo only routine clinical consent. The Zelen “randomized consent” design is preferred by clinicians who are concerned that the informed consent procedure impairs the physician-patient relationship and patient accrual. They may be right, for the

introduction of the Zelen allocation strategy into a flagging breast cancer trial was followed by a 6-fold increase in its recruitment rate<sup>23</sup>. Not surprisingly, some ethicists are concerned about including outcome and other information from control patients who haven't undergone informed consent.

Once trial patients are allocated, their regimens have to be applied, adhered to, and monitored. The next section of this chapter deals with interventions (including non-drug manoeuvres), placebos, compliance, protocol adherence, and compliance. Before we go there, however, we invite you to take a detour and consider an absolute prerequisite for allocating any patient to any trial intervention, uncertainty.

The update of our RRPCE patient-flow diagram appears in Figure 3-6-2.

Figure 3-6-2: RRPCE patient-flow diagram after allocation



## REFERENCES

<sup>1</sup> <http://www.bbc.co.uk/cult/hitchhikers/>

<sup>2</sup> Schulz KF, Grimes DA. Allocation concealment in randomised trials: defending against deciphering. *Lancet*. 2002;359:614-8.

<sup>3</sup> Kunz R, Vist G, Oxman AD. Randomisation to protect against selection bias in healthcare trials (Cochrane Methodology Review). In: *The Cochrane Library*, Issue 4, 2003. Chichester, UK: John Wiley & Sons, Ltd.

<sup>4</sup> Altman D: personal communication to Dave Sackett, 2002.

<sup>5</sup> University Group Diabetes Program Research Group. A study of the effects of hypoglycaemic agents on vascular complications in patients with adult-onset diabetes: II. Mortality results. *Diabetes* 1970;19 (Suppl 2):785-830.

<sup>6</sup> Seltzer H: A summary of criticisms of the findings and conclusions of the University Group Diabetes Program. *Diabetes* 1972;21:976-9.

<sup>7</sup> Cornfield J: The University Group Diabetes Program: A further statistical analysis of the mortality findings. *JAMA* 1971;217:1676-87.

- 
- <sup>8</sup> Taves DR: Minimization: a new method of assigning patients to treatment and control groups. *Clin Pharmacol Ther.* 1974;15:443-53.
- <sup>9</sup> Haynes RB, Sackett DL, Gibson ES, Taylor DW, Hackett BC, Roberts RS, Johnson AL. Improvement of medication compliance in uncontrolled hypertension. *Lancet.* 1976;1:1265-8.
- <sup>10</sup> Schulz KF, Grimes DA. Generation of allocation sequences in randomized trials: chance, not choice. *Lancet* 2002;359:515-9.
- <sup>11</sup> Treasure T, MacRae KD. Minimization: the platinum standard for trials? *BMJ* 1998;317:362-3.
- <sup>12</sup> Peto R, Clinical trial methodology. *Biomedicine Special Issue* 1978;28:24-36.
- <sup>13</sup> Devereaux PJ, Manns BJ, Ghali WA, Quan H, Lacchetti C, Montori VM, Bhandari M, Guyatt GH. Physician interpretations and textbook definitions of blinding terminology in randomized controlled trials. *JAMA* 2001;285:2000-3.
- <sup>14</sup> Schulz KF, Grimes DA. Allocation concealment in randomized trials: defending against deciphering. *Lancet* 2002;359:614-8.
- <sup>15</sup> Moseley JB, O'Malley K, Petersen NJ, Menke TJ, Brody BA, Kuykendall DH, Hollingsworth JC, Aston CM, Wray NP. A controlled trial of arthroscopic surgery for osteoarthritis of the knee. *N Engl J Med* 2002;347:81-8.
- <sup>16</sup> Horng S, Miller FG. Is placebo surgery unethical? *N Engl J Med* 2002;347:137-9.
- <sup>17</sup> Meinert CL. *Clinical Trials; Design, Conduct and Analysis.* Oxford: Oxford University Press 1986. Pp 72-4.
- <sup>18</sup> Friedman LM, Furberg CD, DeMets DL. *Fundamentals of Clinical Trials, 3<sup>rd</sup> edition.* Berlin: Springer, 1998. Pp 72-3.
- <sup>19</sup> Sackett DL. Why randomized controlled trials fail but needn't: 2. Failure to employ physiological statistics, or the only formula a clinical trialists is ever likely to need (or understand!). *CMAJ* 2001;165:1226-37.
- <sup>20</sup> Kahn KS, Daya S, Colins JA, Walter SD. Empirical evidence of bias in infertility research: overestimation of treatment effect in crossover trials using pregnancy as the outcome measure. *Fertil Steril* 1996;65:939-45.
- <sup>21</sup> Pocock SJ: *Clinical Trials A Practical Approach.* Chichester: John Wiley & Sons, 1983. Pp 110-22.
- <sup>22</sup> Zelen M: A new design for randomized clinical trials. *N Engl J Med* 1979;300:1242-5.
- <sup>23</sup> Ellenberg SS: Randomization designs in comparative clinical trials. *N Engl J Med* 1984;310:1404-8.