Contents

- **Why Randomize?**
  - To safeguard against selection bias
  - To try to avoid accidental bias
  - To provide a basis for statistical tests
    - t-tests etc can be justified on basis of the randomization — don't need to appeal to theory of Normal distributions: the tdistribution is a good approximation to the randomization distribution — shewn by Student [W.S. Gossett]

- **Simple randomization**
  - randomization list
    - List assigning next subject to treatment constructed using random numbers
    - Made **before** trial starts
    - Easy to produce using computer package

- **Restricted Randomization**
  - **Blocking**
    - ensures equal treatment numbers at certain equally spaced points in the sequence of patient assignments.
    - next random digit assigns a block of treatments
    - easy to guess next treatment in small blocks
    - large block size *vs* potential imbalance

- **Unequal Collection**
  - may want a treatment used more frequently than others — use blocks of treatments repeated in desired ratios

    e.g. AAABB (permuted), 5!/3!2!=10 possibilities, AAABB, AABAB, …… choose sequence of blocks randomly from list of permutations

- NB may have large number of different permutations — no need to use them all; a subset is adequate
  - 5 treatments equally replicated
  - need list for 200 subjects
  - 120 different blocks of size 5, but only need 40
  - use (say) 10 or 15 of these to construct list — avoids enumerating all 120

Stratified Randomization

- Strata defined by combinations of relevant patient factors
- Prepare separate randomization lists for each strata
- Alternative is to include strata indicators in analysis (e.g. by regression)

- Implementation in **R**
  - key command is **`sample(.)`**

```
> x<- c(0:9)
> x
 [1] 0 1 2 3 4 5 6 7 8 9
> sample(x)
 [1] 6 3 1 7 5 4 9 8 0 2      permutation
> sample(x,4)
[1] 3 1 6 7      subsample without replacement
> sample(x,4,replace=TRUE)
[1] 0 9 0 7 subsample with replacement
> sample(x,20,replace=T)
 [1] 3 8 1 4 0 9 4 7 5 1 6 4 2 3 1 8 3 3 7 0
```



```
> z<-c(rep("A",5),rep("B",5),rep("C",5))
> z
 [1] "A" "A" "A" "A" "A" "B" "B" "B" "B"
 "B" "C" "C" "C" "C" "C"
> sample(z)
 [1] "B" "A" "A" "A" "C" "C" "B" "B" "C"
 "A" "B" "C" "B" "A" "C“
> sample(c(rep("A",4),rep("P",2)))
[1] "A" "A" "P" "A" "P" "A"
```

How can you produce a randomization list of length 24 with blocks of this form??

Minimization

- large number of relevant factors
  - very large number of strata
    - some combinations of factors very rare
- separate randomization lists unrealistic
  - Determine new subjects factor status
  - Count numbers of subjects with those factors on each treatment — allocate to balance up scores (see course notes)

Randomization Software

- A directory of randomisation software is maintained by Martin Bland at:

http://www-users.york.ac.uk/~mb55/guide/randsery.htm

- Downloadable programmes for simple and blocked randomization [some free]
- Easy to programme in **R**
- some commercial software including add-ons for standard packages such as STATA
- links to various commercial *randomization services* used to provide full blinding of trials
- notes on randomization with references etc.

- **Summary and Conclusions**
  - Protects against accidental & selection bias
  - provides a basis for statistical tests (e.g. use of normal and tdistributions)

- simple
  - but may be unbalanced over treatments
- blocked
  - but small blocks may be decoded
- stratified
  - but may require small blocks
- minimization
  - but lessens randomness