JOHNS HOPKINS
UNIVERSITY

Center for Clinical Trials

Department of Biostatistics
Department of Epidemiology
Department of International Health

Department of Medicine
Department of Ophthalmology
Oncology Center

(Wednesday 5:01am) 18 April 2001

Memorandum

To: Center for Clinical Trials faculty and staff

Fr: Curt Meinert

Re: On stratification in trials

Definitions

block *n* - [ME *blok*, fr MF *bloc*, fr MD *blok*; akin to OHG *bloh* block, MIr *blog* fragment] 1. [general] A group, quantity, section, or segment that is considered as a **unit** for some specified purpose, procedure, process, or action; **stratum**. 2. [**trials**] A grouping of **treatments** or of **treatment assignments** administered or to be administered in the order listed; especially a grouping of assignments in a **parallel treatment design** that satisfy the **assignment ratio**. 3. The **experimental unit** or collection of units receiving the treatments represented in a block; a single unit in the case of **crossover designs** and as many units as there are treatment assignments represented in a block in **parallel treatment designs**. syn (trials): **treatment assignment block** rt: **blocked treatment assignment**, **blocked randomization** ***Usage note***: Not to be confused with **stratum** (although often so confused, especially in lay usage, as suggested by defn 1); see note for **stratification**. The **block size** may be the same for all blocks, eg, as required for a **crossover treatment design** and as arising in **parallel treatment designs** involving blocks of the same size. The size will be equal to the number of **study treatments** in the case of **complete crossover treatment designs**. The size for parallel treatment designs will be the sum of the numbers represented in the **assignment ratio** (eg, 2 for a design involving two study treatments and a **uniform treatment assignment ratio** and 15 for a design involving 6 study treatments and an **assignment ratio** of 1:1:1:1:1:2.5) or multiples thereof. The usual strategy in parallel treatment designs is to have a mix of blocks of different sizes, themselves randomly ordered, with all blocks being some multiple of the smallest possible block size. The purpose of blocking is to ensure **balance** in the mix of the **treatment groups** with regard to time of enrollment. The blocking helps to eliminate the risk of differential **secular trends** in the make-up of the treatment groups. Time-related shifts in the nature of persons **enrolled** over the course of a trial can be a **confounding variable** for **treatment comparisons**, in the absence of blocking, if the mix of persons represented changes over time and is different by treatment group.

blocked randomization *n* - 1. **Blocked treatment assignments** in which the assignments are the result of some **randomization** process. 2. **Randomization** constrained by **blocking**. rt: **restricted treatment assignment**

Room 5010 C 615 North Wolfe Street C Baltimore, Maryland 21205 C 410 955-8175 C Fax: 410 955-0932
E-mail: cctmail@jhsph.edu C Web site: http://www.jhsph.edu/Research/Centers/CCT

quota *n* - [ML, fr L *quota pars* how great a part] 1. A proportional part or share. 2. The number or amount representing or constituting a proportional share or limit, as in a **recruitment quota**.

quota requirement *n* - [**trials**] A **requirement** imposed to ensure a specified mix in the **study population** with regard to one or more **demographic variables**, disease state variables, enrollment site, or other variables observed at or prior to **enrollment**; stated in numbers required, proportions required, or as minimums required; eg, in regard to **gender**, 50 males and 50 females, 50% female, or at least 50% female. rt: recruitment quota, **sample size requirement** ***Usage note***: See **quotification**.

quotification *n* - The act or process of imposing a **quota requirement** on the mix of persons **enrolled** into a **trial**. rt: **quota requirement**, **stratification** ***Usage note***: Not to be confused with **stratification**. The purpose of stratification is to ensure that the different **treatment groups** all have the same proportionate mix of people with regard to the **stratification variable**(s). The purpose of quotification is to ensure a **study population** having a specified mix with regard to the variables used for quotification.

restricted randomization *n* - **Randomization** involving restrictions, such as in **blocked randomization**; not **complete randomization**. ant: **complete randomization**, unrestricted randomization

stratified randomization *n* - A **treatment assignment process** using **stratified random treatment assignment**. ***Usage note***: See **stratified-blocked randomization**.

stratified-blocked randomization *n* - **Stratified randomization** constrained by **blocking** within **strata**. ***Usage note***: Largely synonymous with **stratified randomization** because virtually all occurrences of the term are in relation to schemes involving blocking. **Stratification**, in the absence of blocking (or some other restriction having a similar effect) to force assignments to meet the specified **treatment assignment ratio** at points over the **enrollment process**, is not likely to be effective in achieving its desired ends.

treatment assignment ratio *n* - 1. The **ratio** of the number of **treatment units** (usually persons) to be or actually in one **treatment group** relative to another, eg, a ratio of 1:2 for a particular **test-treated group** relative to the **control-treated group**. 2. The ratio of one group relative to all others, eg, a ratio of 1:1:1:1:1:2.5 for a trial with five **test treatments** and a **control-assigned group** that is or is to be 2.5 times larger than any of the **test-assigned groups**.

stratified randomization *n* - A **treatment assignment process** using **stratified random treatment assignment**.

stratified random treatment assignment *n* - **Random treatment assignment** within defined **assignment strata**.

stratification *n* - 1. Broadly, the act or process of **stratifying** (defn 2 or 3). 2. An active ongoing process of **stratifying**, as in the sense of defn 2, as in placing **patients** into **strata** as they arrive at a **clinic** as a prelude to **enrollment** and **randomization** to **treatment** in a **trial**.

3. The act or process of classifying **treatment units** or **observations** into strata after **enrollment** for a **subgroup analysis**; **post-stratification**. rt: **classification**, **quotification** ***Usage note***: Stratification is done as a means of **controlling** sources of **variation** related to or assumed to be related to the **outcome**. Stratification (defn 2) and **blocking** in the **treatment assignment** process serve different purposes. Blocking is imposed as a means of ensuring that the **assignment ratio** will be satisfied or nearly satisfied; stratification is done to ensure the comparability of the treatment groups with regard to the variable(s) used in stratification. There is confusion regarding the meaning and impact of stratification on the design and operation of a trial. Often the act of stratification is taken as evidence of the need to perform **treatment comparisons** within the various strata represented in the stratification. Although that may be desirable, such comparisons are not necessary. Valid comparisons of the **treatment groups** can be performed by pooling across strata. As a rule, the mix of persons enrolled into a trial is determined by the mix of persons seen and ultimately judged eligible for **enrollment**. Hence, the numbers to be represented in the various strata will be **variables** having values known only after completion of enrollment. The imposition of a **sample size requirement** for one or more of the strata (see **recruitment quota**), in addition to one for the trial, extends the time required for **recruitment** and should not be imposed unless there are valid scientific or practical reasons for doing so. Confusion also arises from use of the term **stratification** in two distinctly different contexts, as suggested in defns 2 and 3 above. Use **post-stratification** for uses in the sense of defn 3, especially when in settings, such as trials, where both forms of stratification are used.

stratification variable *n* - 1. A **variable** used to classify **treatment units** into **strata** in relation to **treatment assignment**. 2. A variable used to classify **observation units** into strata in relation to **data analysis**.

Introduction

Stratification is done to ensure balance in the distribution of a variable (or set of variables) across treatment groups in a trial. The balance is achieved by restricting the randomization so as to ensure that the assignment ratio is satisfied within strata. Typically, because the numbers to be enrolled by strata are unknown, randomization is done in blocks over the course of enrollment.

Stratification is not to be confused with quotification (see memo of 11 May 2001). Stratification is done to ensure the same mix across strata, whereas quotification is done to ensure a specified mix in the trial.

Stratification is an active process that takes place in conjunction with randomization. It is not to be confused with subgrouping, as discussed in a memo on subgroup analyses (11 March 2001).

Stratification is used differently by epidemiologists and trialists. The epidemiologist uses the term in relation to sorting done for subgroup analyses, whereas the trialist reserves use for sorting done in relation to randomization. The trialist uses *subgrouping* or *post-stratification* to refer to sorting done in relation to data analyses.

The effect of the stratification and blocking is to "control" for the stratification variable in the assignment process. For example, suppose the variable is gender, then the effect of the

stratification and blocking is to ensure the same gender mix across treatment groups. If the gender mix for the trial is 25:75, that mix (within the limits of blocking) will be maintained across treatment groups in the trial. If the mix is 75:25 that mix will be maintained across treatment groups.

Stratification does not eliminate need for adjustment for differences in the baseline composition of the treatment groups. The only variables "controlled" are those used for stratification. Stratification does nothing to control variation of other baseline variables. In fact, the "control" effected may not be sufficient to obviate need for adjustment even for the stratification variables. The variation in the mix of people across treatment groups can be widely variant over the course of enrollment if blocks are large.

Restricted randomization

Restricted randomization via blocking, in the context of trials, is done to ensure that the assignment ratio specified in the design is met. The advantage of blocking is that it reduces the risk of confounded treatment assignments – a condition where all or a disproportionately large number of persons are assigned to the same treatment group.

To illustrate the risk of confounding consider a trial involving two treatment groups (A and B), a planned sample size of 100, and an assignment ratio of 1:1. It is possible (even if unlikely), in the absence of blocking, to end up with all 100 assignments to same treatment (1st line; Panel A). The maximum difference in numbers assigned with blocking (column F) is half the block size.

The assignment ratios corresponding to values recorded in Panel A are given in Panel B of the table. Column F corresponds to the desired assignment ratio (ie, 1.00) minus the smallest observable ratio at the four points in enrollment.

Maximum allowable observed difference in assignments (Panel A) and the corresponding observed assignment ratio (Panel B) at different points in enrollment

A	B	C	D	E	F
	Number enrolled				
Blk size	25	50	75	100	Max diff
A. Maximum observed difference in assignments (A to B or B to A)					
No blocking	0 25	0 50	0 75	0 100	100
200	0 25	0 50	0 75	0 100	100
100	0 25	0 50	25 75	50 50	50
50	0 25	25 25	25 50	50 50	25
30	10 15	20 30	30 45	45 55	15
20	10 15	20 30	35 40	50 50	10
10	10 15	25 25	35 40	50 50	5
8	12 13	24 26	36 39	48 52	4
4	12 13	24 26	37 38	50 50	2
2	12 13	25 25	37 38	50 50	1
B. Maximum observable assignment ratio					
No blocking	0	0	0	0	1.00
200	0	0	0	0	1.00
100	0	0	0.33	1.00	1.00
50	0	1.00	0.50	1.00	1.00
30	0.67	0.67	0.67	0.82	0.33
20	0.67	0.67	0.67	1.00	0.33
10	0.67	1.00	0.88	1.00	0.33
8	0.92	0.92	0.92	0.92	0.08
4	0.92	0.92	0.97	1.00	0.08
2	0.92	1.00	0.97	1.00	0.08

The smaller the block size, the smaller the departures from the desired assignment ratio. But the more restrictive the scheme the greater the risk of treatment-related selection bias if the blocking scheme becomes known to clinic personnel. For example, with blocks of size 2, clinic personnel would be able to predict half the assignments in advance of release.

Choice of stratification variables

The trialist is responsible for selecting the variables to be used for stratification, if any. Choices include:

- Demographic variable
  - Gender
  - Ethnic origin
  - Age
  - Education
  - Socioeconomic status
- History variable
  - Family history
  - Medical history
  - Treatment history
- Geographic variable
  - Area of residence
  - Distance from clinic
  - Clinic (in multicenter trials)
- Option variable
  - Treatment menu
  - Physician\surgeon
- Physiological variable
  - Blood pressure
  - Cholesterol
  - Hematocrit
  - Height
  - Body weight
  - Body mass index
- Reading variable
  - Biopsy
  - ECG
  - Fundus photograph
  - Angiogram

The stratification variable must be independent of treatment assignment – a virtual certainty because the variable has to be observed prior to randomization. Independence is assured if treatment assignments are concealed in advance of issue.

In addition, the variable should be reliable, ie, capable of giving the same value on replication. Use of a variable lacking reliability means that the stratification is prone to classification errors. The value of stratification for variance control diminishes with increasing rates of misclassification.

A third desirable (but not essential) property of stratification variables is amenability to "on the spot" observation or measurement. Use of variables that have to be assessed by sending records or specimens to central reading facility or laboratory complicates the randomization because assignments cannot be issued until the reading or determination has been made and is back at the clinic.

Reasons for stratification

Quantitative interaction

A quantitative interaction is one in which the direction or sign of the relationship is the same but the magnitude of the difference depends on the value assumed by the stratification variable; eg, a treatment effect that is different but of the same sign for males versus females. The estimate of treatment effect depends on the value of the interacting variable. The best estimate of effect differential is when the mix of assignments is the same across strata. The best way to assure that that wil be the case is by stratification on the interacting variable and by blocking within strata.

Note that the reference is to *quantitative* interaction. The proper strategy in the presence of an expected *qualitative* interaction is to exclude persons not expected to benefit from treatment rather than to stratify.

Designed subgroup comparisons

A designed subgroup comparison is one specified in design documents of the trial and, typically, for which there is a required sample size. The trialist stratifies on the subgrouping variable for the same reason as cited above in regard to suspected quantitative interactions.

Variance control

A major reason for stratification is for variance control. But to be useful in this regard the variable has to be predictive of outcome. Absent that there is no gain.

The reality is that many more things are assumed to be predicative than are in fact predicative. Hence, many variables chosen because of their presumed predictive value turn out to have little or no predictive power.

The other reality is that the increase in precision achieved by stratification, even when variables are predictive, diminishes with increasing sample size. The gain is minimal for sample sizes in excess of 50 per treatment group.

Practical and logistical simplification

Another reason for stratification has to do with practical and logistical issues. For example, one would stratify on gender if the menu of treatments for males and females is different. Similarly, one is inclined to stratify on clinic in multicenter drug trials to "normalize" drug requirements across clinics. The stratification allows the coordinating center to project needs on a clinic by clinic basis, thereby simplifying supply logistics.

Protection against confounding within subgroups

Another reason for stratification is to avoid confounding. Confounding occurs if an entry characteristic is differentially distributed by treatment group and if that characteristic is related to outcome. Confounding is avoided if the distribution of that variable is the same across treatment groups over the course of enrollment. One way to ensure that that is the case is by stratified-blocked randomization, but there are limits as discussed below.

Political considerations

There are occasions when it is simply politically expedient to stratify even though the scientific basis for stratification is lacking. For example, trialists may opt to stratify on gender, age, or race merely because those variables have been touted to be important in moderating treatment effect.

On the limits of stratification

The purpose of blocking and stratification is to restrict randomization so "force" assignments to meet the desired assignment ratio at different points over the course of enrollment. But there are no free lunches. "Too much" stratification can be the operational equivalent of none in the "forcing" process.

To illustrate, suppose a single binary type stratification variable (eg, gender), a 1:1 assignment ratio, and blocks of size 8. The maximum permissible departure from the desired assignment ratio of 1:1 without stratification is when a block is half filled and when all assignments in the block up to that point are to the same treatment group. The maximum is 8 with the stratification; occurring when, by chance, the block in both strata is half filled and where the assignments are all to the same treatment group in both strata. The maximum is 16 with two binary type stratification variables, 32 with three such variables, etc.

The potential for complete confounding in the trial is not eliminated until enrollment exceeds (blk size/# of trts)x(no. of strata), eg, 64 (ie, $(8/2)x2^4$) with for 4 binary type variables. The risk can be reduced with smaller block sizes. For example, halved with blocks of size 4, and halved again with blocks of size 2, but the smaller the block size the more likely it is that clinic personnel may be able to predict assignments and, hence, increase the risk of selection bias.

Recommendations to trialists regarding stratification

1. Limit choice to variables known or plausibly believed to influence outcome

2. Limit to a small number of variables; not more than 2 or 3

3. Limit choice to variables that are evaluated by direct interview or observation such as medical history, gender, age, or body weight; variables based on laboratory results or readings of pathology specimens may cause delays in randomization

4. Shy away from stratification when the treatment groups are large (except for clinic in

multicenter)

Distribution

via local e-mail

(See names8.wpd list)

CLM Musing file

Chronologic file