



Computer Randomization Algorithms: Drug Development Pipeline

Natalie LePera, STA514 – West Chester University



Computerized Randomization in the Drug Development Pipeline



Metabolomics / Proteomics: Mass Spectrometry Plate Randomization

Converts biological samples into digital / data
Pre-Clinical studies (animal models) for identification of target protein/gene
Randomization reduces machine drift & batch effects

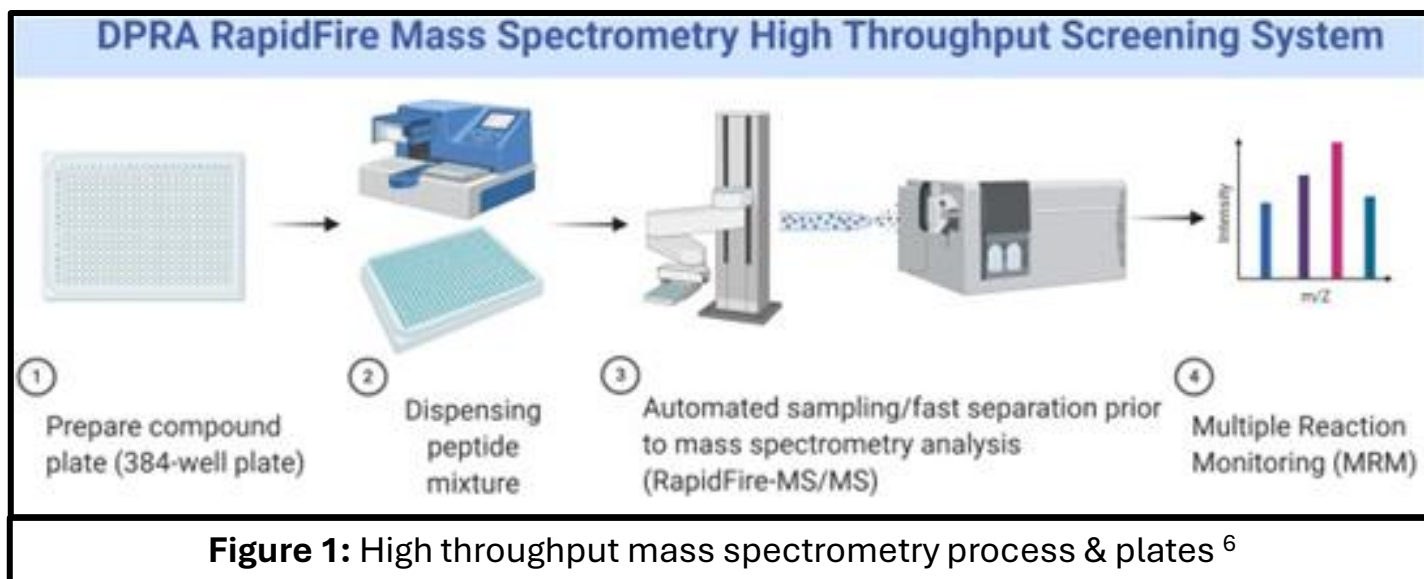


Double Blind Clinical Trials: Interactive Web Randomization Service (IWRS)

Randomizes patients into treatment / placebo group(s)
Uses stratification as pseudo matching between experimental groups (age, gender, diabetes, etc.)
Customizable to trial specifics

Mass Spectrometry (MS): Intro

- Biological samples collected to assess proteins, metabolites, and/or nucleic acid content ⁴
- Samples loaded into wells on plates for bulk analysis
 - Automated injection processes programmed by block randomization software ⁴
 - Commonly >350 wells per plate ³



Common Problems

- Machine drift caused by contamination, temperature fluctuations, aging components, etc.
 - Signal can “drift” up to 25% in 12hrs ³
- Samples often received by lab **not in** chronological order but collated
 - Requires documentation and traceability of all randomization ⁴

MS Randomization: Inter-Batch ⁴

Goal: Randomizes injection order of samples into batches to reduce signal drift

- **Batches:** usually full plates but can be mini squares within the plate depending on lab process

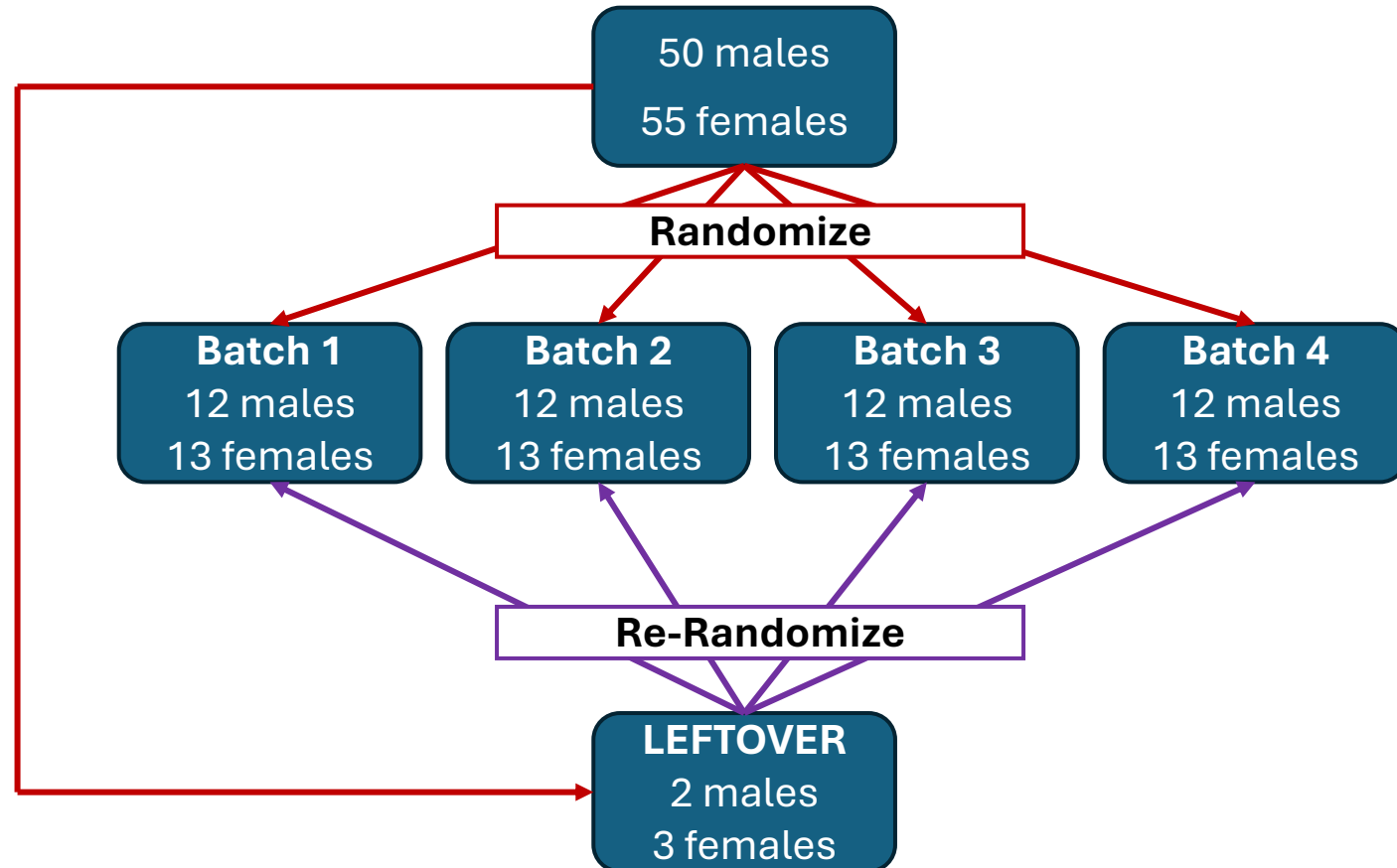
Stratified randomization utilized

- Independent of any study randomization
- Strata pre-selected – commonly **potentially confounding variables**

Reduces batch interference by splitting samples into batch ensuring even distribution of identified prognostic factors

Re-randomizes “leftover” tail samples that do not cleanly fit into a group

EXAMPLE:



MS Randomization: Intra-Batch⁴

Block Randomization Algorithm

- Dimension (dim) size set as total number of sample groups (prognostic factor from Inter-batch)
 - Example: 4 dimensions = Batch 1, Batch 2, Batch 3, Batch 4
- The size of the sample “block” is selected as a multiple of the dim size
 - Block size for 2 dimensions = 2, 4, 6 etc.
 - Minimum Recommended block size = $2 * \sum(\text{tx ratio parts})$
 - Example: 2 treatments in 1:1 ratio = $2(1+1)$ = Minimum block: 4
- Write out all possible “block” combinations of samples that can be placed in the assigned block size & number
- Generate random number sequence using range of 1 – number of combinations
 - Example: Range_(dim = 2, block = 4) = (1,6) | sequence: 1,5,2,2,3,6,4,5,5
- Populate sequence using combination blocks as listed in the random sequence from step 4

| Dim Size | Block Size | Block Permutation | Generate Random Sequence | Block Replacement Result |
|----------|------------|----------------------------------|--------------------------|------------------------------------|
| 2 | 2 | AB BA | 1,2,1,2,2,1,2,1 | AB BA AB BA BA AB BA AB |
| 3 | 3 | ABC ACB CAB BAC BCA CBA | 1,2,4,2,6,5,5,3 | ABC ACB BAC ACB CBA BCA BCA CAB |
| 2 | 4 | ABAB BBAA AABB BABA BAAB ABBA | 3,6,2,5,1,6,4 | AABB ABBA BBAA BAAB ABAB ABBA BABA |

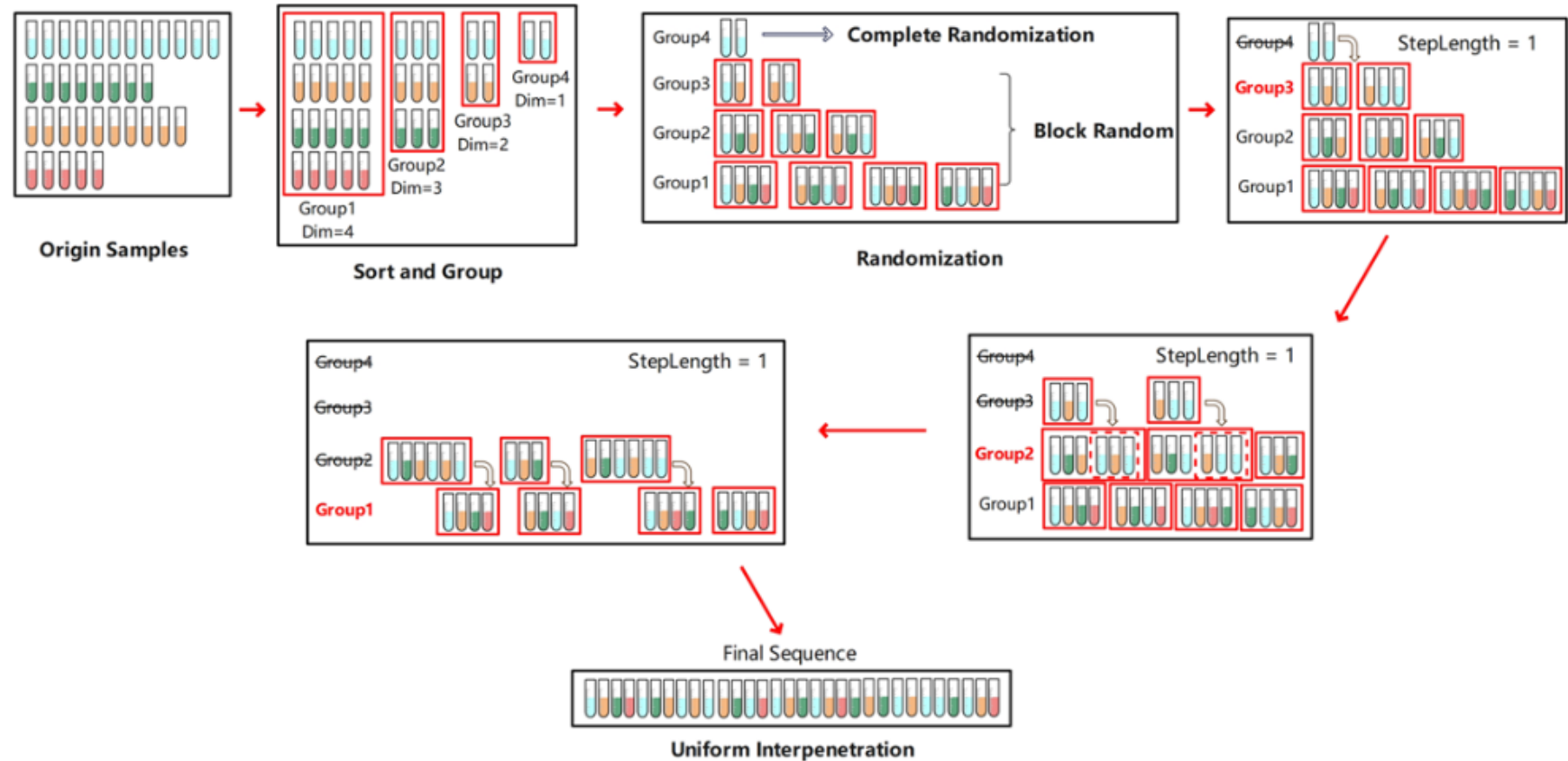
Figure 2: Visual representation of block randomization of samples from treatment groups A and B⁴

Goals:

- Reduces impact of machine drift over multiple sample runs
- Allows for QC samples to be collected from the same wells each run, as content of wells randomized
 - Reduces impact of manual fatigue after repeated sample filling if no automated fill

MS Randomization: Intra-Batch Uniform vs Non-Uniform Sample Distribution⁴

Non-Uniformly Distributed Samples Adaptation



Figures 3 : Randomization of non-uniformly distributed samples & uniformly distributed samples⁴

+

•

○

Clinical Trial Randomization: Interactive Web Randomization Service (IWRS)

Centralized web applications allow for automated randomization into clinical trials

- Allow for 24/7 emergency blind breaking
- Ensures investigative team is blinded to treatment assignment
- Allows for blocking & stratification programming for prognostic factors ²
- Removes selection bias and prevents accidental bias
- Allows for randomization into multiple treatment groups + placebo group
- Allows for unequal allocation to treatment groups (3:1:1:1) ⁵

Common approaches:

- Pocock-Simon's Imbalance Score
- The Symmetric KLD Index
- Two-Way Minimization Method ⁵

IWRS: Pocock-Simon's Imbalance Score ⁵

$$G_i = \sum_{j=1}^J \sum_{m=r_j} n_{ijm}$$

where

the next patient entering has levels r_1, r_2, \dots, r_j on prognostic factors (j) 1, 2, ..., J

- Continuous prognosis factors are categorized to calculate treatment imbalance
 - Includes an initial simple randomization for **a programmable number initial patients randomized** before beginning imbalance calculation
- Calculates marginal treatment total (G_i) if the new incoming patient is assigned to treatment group (i)
 - Relies on defined prognosis factors (J), levels of factor J (M_j), and treatment arm (i)
- Totals for each treatment (i) are ranked to assign weights for randomization
 - Least G_i total given highest weight
 - Lower G_i total indicates the treatment that causes least imbalance between groups

+

•

○

IWRS: Systemic KLD Index ⁵

$$d_i = \frac{1}{2} \sum_{j=1}^{j'} \left[\left\{ (\bar{x}_{ij} - \bar{x}_{i'j})^2 + (V_{ij} + V_{i'j})^2 \right\} * \left(\frac{1}{V_{ij}} + \frac{1}{V_{i'j}} \right) - 2j' \right] + \sum_{j=j'+1}^J \sum_{m=1}^{M_j} (p_{ijm} - p_{i'jm}) * \log \frac{p_{ijm}}{p_{i'jm}}$$

where $\bar{x}_{ij} = \frac{1}{n_i} \sum_{k=1}^{n_i} x_{ijk}$
 and $V_{ij} = \frac{1}{n_i - 1} \sum_{k=1}^{n_i} (x_{ijk} - \bar{x}_{ij})^2$

- Approach limited to **two treatment arms**
- Includes an initial simple randomization for the **first 4 patients**
- Assumes **multivariate distribution** of continuous prognosis factors but works with normal distributed continuous & categorical factors
- Calculates difference in distribution of prognosis factors between 2 treatments i and i' (d_i)
 - Uses \bar{x}_{ij} as mean of prognostic factor value for pts in treatment i with prognostic factor j
 - Uses p_{ijm} as the proportion of subjects assigned to level m of prognostic factor j
 - Relies on defined prognosis factors (J), levels of factor J (M_j), and treatment arm (i)
- Compares treatment arms with each other and with placebo to calculate d_i values
 - $d_{i \text{ tx1 vs tx2}}$, $d_{i \text{ tx2 vs tx1}}$, $d_{i \text{ tx1 vs placebo}}$, $d_{i \text{ placebo vs tx1}}$, $d_{i \text{ tx2 vs placebo}}$, & $d_{i \text{ placebo vs tx2}}$
- Totals for each treatment (i) are ranked to assign weights for randomization
 - Least d_i total given highest weight
 - Lower d_i total indicates the treatment that causes least imbalance between groups

IWRS: Two-Way Minimization⁵

Imbalance in total number of subjects:

$$\delta = |\eta_T - \eta_C|$$

Imbalance in distribution of prognostic factors:

$$D = \sum_{j=1}^J \left(\frac{\sum_{m=1}^{M_j} |p_{Tjm} - p_{Cjm}|}{M_j} \right)$$

- Approach limited to **two treatment arms**
- Includes an initial simple randomization until **1 subject is randomized to each group**
- Calculates both imbalance in total subject number **and** in distribution of prognostic factors
 - Uses n_T as total allocated to treatment & n_C as total allocated to control (placebo)
 - Uses p_{Tjm} as proportion in treatment with prognostic factors j at level m & p_{Cjm} as proportion in control with prognostic factors j at level m (placebo)
- Calculate probability the new patient is allocated to minimize imbalance in patient number (P), and the probability the patient is allocated to minimize prognostic factor distribution ($1-P$)
 - P calculated using $P = 1 - 0.95^\delta$





Reference List

1. Daghlas I, Gill D. Mendelian randomization as a tool to inform drug development using human genetics. *Cambridge Prisms: Precision Medicine*. 2023;1:1-12. doi:10.1017/pcm.2023.5
2. Jatinder Singh. Randomization and online databases for clinical trials. *Journal of Pharmacology & Pharmacotherapeutics*. 2014;5(2):173-174. doi:10.4103/0976-500X.130155
3. Jiang F, Liu Q, Li Q, et al. Signal Drift in Liquid Chromatography Tandem Mass Spectrometry and Its Internal Standard Calibration Strategy for Quantitative Analysis. *Analytical Chemistry*. 2020/06/02 2020;92(11):7690-7698. doi:10.1021/acs.analchem.0c00633
4. Lu M, Jiang H, Wang R, An S, Wang J, Yu C. Injection design: web service of plate design with optimized stratified block randomization for modern GC/LC-MS-based sample preparation. *BMC Bioinformatics*. 2023;24(1):1-16. doi:10.1186/s12859-023-05598-1
5. Xiao L, Huang Q, Yank V, Ma J. An Easily Accessible Web-Based Minimization Random Allocation System for Clinical Trials. *Journal of Medical Internet Research*. 2013;15(7):1. doi:10.2196/jmir.2392
6. Wei Z, Fang Y, Gosztyla ML, et al. A direct peptide reactivity assay using a high-throughput mass spectrometry screening platform for detection of skin sensitizers. *Toxicology Letters*. 2021/03/01/ 2021;338:67-77. doi:https://doi.org/10.1016/j.toxlet.2020.12.002