

· 学术探讨 ·

基于临床研究数据模拟比较简单随机、动态均衡随机及最小化法随机的均衡性

李庚¹, 温泽淮^{1,2} (1. 广州中医药大学第二附属医院/广东省中医院, 广东 广州 510120; 2. 广州中医药大学 DME 中心, 广东 广州 510405)

摘要: **目的** 探讨简单随机、动态均衡随机、最小化法随机 3 种不同随机方法的均衡性。**方法** 利用真实临床研究数据进行模拟测试, 比较 3 种不同随机化方法分配结果的组间例数均衡性和非处理因素组间均衡性。**结果** 在样本量分别为 20、100、200、500 和 1 000 条件下, 动态均衡随机比最小化法和简单随机拥有更好的均衡性, 具体表现在组间例数绝对差值、平均边际不平衡值和最大边际不平衡值; 而最小化法的均衡性又比简单随机好。此外, 动态均衡随机拥有比最小化法更好的随机性。**结论** 动态均衡随机和最小化法 2 种动态随机化方法具有良好的均衡性, 在小样本、非处理因素较多时比较适用, 尤其是当非处理因素重要程度不同时可以考虑采用动态均衡随机化方法。

关键词: 简单随机; 动态均衡随机; 最小化法; 均衡性; 模拟; 临床研究

中图分类号: R969; R285.6; R195.1 **文献标志码:** A **文章编号:** 1003-9783(2021)06-0894-05

doi: 10.19378/j.issn.1003-9783.2021.06.021

Comparison of Simple Randomization, Dynamic Balancing Randomization and Minimization: A Simulation Study Based on Real Clinical Trial Data

LI Geng¹, WEN Zehuai^{1, 2} (1. The Second Affiliated Hospital of Guangzhou University of Chinese Medicine, Guangdong Provincial Hospital of Chinese Medicine, Guangzhou 510120 Guangdong, China; 2. Center of DME of Guangzhou University of Chinese Medicine, Guangzhou 510120 Guangdong, China)

Abstract: **Objective** To explore the balance of simple randomization, dynamic balancing randomization and minimization. **Methods** Real clinical research data was used to conduct simulation tests to explore the balance of three different randomization methods, the equilibrium of cases between groups and equilibrium between groups of non-treatment factors were taken for comparison. **Results** Dynamic balancing randomization consistently produced the best balance in various sample sizes (20, 100, 200, 500 and 1 000) in all three methods, which is specifically showed in the differences of number between groups, average marginal imbalance and maximum marginal imbalance; and minimization method had a better balance than simple randomization. It is also found that dynamic balancing randomization had a better randomness than minimization method. **Conclusion** This study demonstrates that dynamic balancing randomization and minimization method outperforms simple randomization with regard to achieving balance. They are suitable for trials with small sample size and multiple prognostic factors. When the importance of non-treatment factors is different, dynamic balancing randomization method can be considered.

Keywords: Simple randomization; dynamic balancing randomization; minimization; balance; simulation; clinical trials

收稿日期: 2020-07-15

作者简介: 李庚, 女, 博士研究生, 助理研究员, 研究方向: 临床研究方法学。Email: ligeng@gzucm.edu.cn。通信作者: 温泽淮, 男, 博士, 研究员, 博士研究生导师, 研究方向: 临床研究方法学。Email: wenzh@gzucm.edu.cn。

基金项目: 广东省中医院中医药科学技术研究专项课题(YN2018QL10)。

1935 年 Fisher^[1]首次提出将随机化(Randomization)应用于实验设计,经过近 80 年的应用与发展,随机化已成为临床研究设计遵循的基本原则之一。随机化原则包括随机抽样和随机分配 2 种形式。随机分配是指纳入研究的合格对象有相同或/和固定的概率被分配到研究的任何一个组别,从而一方面使已知的和未知的影响因素在各处理组间达到均衡,另一方面避免研究者或受试者主观意愿的干扰而达到减少选择性偏倚和增加统计推断效力的目的^[2-5]。

在临床试验中,除了要考虑将要研究的处理因素外,经常还存在很多非处理因素,比如性别、年龄、体质量等可能会对试验结果产生影响的因素。通过随机分配可以控制非处理因素在比较组间的均衡^[6]。因此,组间均衡性是评价随机分配方法需要考量的内容之一。均衡的分配结果不但要保证各处理组中纳入的例数是相等的,还要保证重要的非处理因素在比较组间的分布是均衡的^[7-8]。最小化法随机分组的均衡能力已得到一些学者的认可^[9],但也有研究认为动态均衡随机化的均衡能力甚至高于最小化法^[8]。本研究拟以某一临床研究的数据为例,探索不同方法分配结果的均衡性。

1 研究对象、目的与方法

1.1 研究对象 本研究数据来源于中药辨证治疗流行性及普通感冒的多中心随机双盲安慰剂对照试验,旨在评估中药辨证治疗流行性及普通感冒的疗效。该试验共纳入临床确诊的单纯型流行性感冒及普通感冒患者 1 381 例,其可能影响结局的非处理因素包括性别(男、女)、年龄(分为 18~24 岁和 24 岁以上 2 个年龄段)、中医证型(包括外感风热证、外感风寒证和外感夹湿证 3 个证型)和分中心(包括广东省中医院总院、广东省中医院二沙分院、广东省中医院芳村分院和广东省中医院大学城分院 4 个分中心)。表 1 是该试验基线特征中各非处理因素的频率分布情况。

1.2 研究目的 利用“1.1”项试验数据进行事后验证,以比较不同样本量(20、100、200、500、1 000)情况下,3 种随机分配方法(简单随机、动态均衡随机和最小化法)的组间例数均衡性(处理组间总例数的绝对差值)和非处理因素的组间均衡性。

1.3 方法及基本原理

1.3.1 简单随机 简单随机又称完全随机化,是利用随机数字表、计算机程序等方法产生随机数字序列,在此基础上安排对象的分组^[11],它不受患者基线

表 1 本研究数据的非处理因素及其频率分布[例(%)]

Table 1 Distributions of the covariates of our study[n(%)]

非处理因素	频率分布
性别	
男	528(38.2)
女	853(61.8)
年龄	
18~24 岁	707(51.2)
>24 岁	674(48.8)
证型	
外感风热证	479(34.7)
外感风寒证	460(33.3)
外感夹湿证	442(32.0)
分中心	
广东省中医院总院	430(31.1)
广东省中医院二沙分院	404(29.3)
广东省中医院芳村分院	149(10.8)
广东省中医院大学城分院	398(28.8)

特征的影响。该操作简单方便且具有高不可预测性^[12]。然而,在样本量较小的情况下,该方法容易造成组间例数不均衡^[9-10,12]。

1.3.2 动态均衡随机 1993 年 Signorini 等^[13]首先提出动态均衡随机化。动态均衡随机化是一个逐步调整的过程。根据试验情况确定多个非处理因素及其重要程度分级,根据其重要程度,逐个计算各因素中组间例数的差值并与所允许最大差值 N_j 比较,给予确定性或随机性分配,逐步调整各组的不均衡情况^[8]。

本研究考虑的非处理因素有:年龄、性别、证型、分中心,这些非处理因素优先顺序是:年龄(其组间允许最大人数差 N_j 为 2)、性别(N_j 为 4)、证型(N_j 为 2)、分中心(N_j 为 2)。随机分配时,先考察优先级别高的非处理因素,若非处理因素内处理组间例数的差值超过或等于 N_j ,则将受试者分入例数较少的组;如未超过,则考虑下一个非处理因素;如果所有的非处理因素亚组的例数差值都未超过 N_j 值,则考虑组间的例数差有没有超过 2,若仍未超过则随机给予处理分配。

1.3.3 最小化法 1974 年 Taves^[14]在偏性掷币(Biased coin)方法的基础上提出最小化法,Pocock 等^[15]在其基础上进行了完善。最小化法的基本原理是通过计算在各组间已经入组受试者的非处理因素(特别是基线资料)的不均衡性,以一定的概率确定下一例受试者的分配,以此来控制对结果有较大影响的非处理因素在组间分布的不均衡性^[7,16]。Pocock 等的最小化法通过 3 个因素确定新病例的组别:(1)各非处理因素在各个处理组间的差别(D),一般可以使用全距或者方差来确定 D 值;(2)各处理组间总体不均衡性

(G); (3)新病例分配到目标组的概率(P)。分组过程中,第1例受试者完全随机分配,随后的受试者采用最小化法,但是如果各处理组间总体不平衡性 G 相等,则采用完全随机化分配^[17]。本研究中,第1例受试者采用简单随机化分配,此后的受试者采用最小化法随机,拟考虑的非处理因素见表1, P 值采用0.8。

1.4 评测指标及统计分析方法

1.4.1 评测指标 (1)组间例数绝对差值 $N^{[7]}$:计算各处理组间例数绝对差值的均数,第25、50、75百分位数以及极大值,以比较不同随机分配方法的组间均衡性。

(2)不平衡值(Imbalance score, B)^[18]:计算非处理因素在不同处理组间(1,2)不平衡值(B)大小。不平衡值(B)的计算方式如下: $B = \sum_{i=1}^C W_i (\bar{x}_{1i} - \bar{x}_{2i})^2$,其中, i 代表第 i 个非处理因素($i=1, 2, \dots, C$), $\bar{x}_{1i}$ 和 $\bar{x}_{2i}$ 是第 i 个非处理因素在组别1和组别2中的平均值, W_i 为第 i 个非处理因素的权重。

(3)边际不平衡性(Marginal imbalance, $\bar{B}_M$, $MaxB_M$)^[19]:设定 b_{Mij} 是第 i 个非处理因素的第 j 个水平的不同组别受试者人数的边际不平衡。其计算公式如下: $b_{Mij} = \frac{|n_{i1} - n_{i2}|}{\sum_{k=1}^2 n_{ijk}}$,其中, n_{ijk} 为第 K 组在非处理因素 i 的第 j 水平上的人数, $\bar{B}_M = \bar{b}_{Mij}$, $MaxB_M = \max_{ij} \{b_{Mij}\}$ 。 $MaxB_M$ 能够很好地反映某个处理因素的大多数水平非常均衡但少数水平却高度不平衡的情况。

(4)不平衡检验的显著性(I_{pi})^[7,18]:通过非处理因素与组别的列联表卡方检验来评价非处理因素的均衡情况。 I_{pi} 表示第 i 个非处理因素的检验显著性,当 $P < 0.05$ 时, I_{pi} 为1;当 $P \geq 0.05$ 时, $I_{pi}=0$ 。

(5)完全随机分配次数^[8]:在不同样本量下,最小化方法和动态均衡随机方法在分配过程中完全随机分配处理的次数。

(6)基线症状的组间均衡性:选取分中心为广东省中医院芳村分院的受试者或随机选取40例受试者,进行动态均衡随机、最小化法随机和简单化随机,并与研究所用的分层区组随机比较基线时受试者症状(鼻塞、咽喉痛、咳嗽、肌肉酸痛、疲劳、头痛)的组间均衡性。

1.4.2 统计分析方法及实现 本研究涉及的统计学检验为双侧检验,显著性水平为 $\alpha = 0.05$ 。根据不同的数据类型,分别采用不同的统计方法,定性资料采用卡方检验,定量资料采用 t 检验。统计分析应用PASW 18.0软件完成,模拟随机化分配采用SQL Sever编程语言实现。

2 结果

2.1 组间均衡性比较 表2、表3显示了不同随机化方法、不同样本量下组间例数绝对差值 N 和边际不平衡性($\bar{B}_M$, $MaxB_M$)分布。结果表明,3种随机化方法中,动态均衡随机化的组间例数绝对差值 N 的均数和四分位数均最小,而简单随机化 N 的均数和四分位数最大。边际不平衡性 $\bar{B}_M$ 和 $MaxB_M$ 结果显示,动态均衡随机化的边际不平衡性最小,最小化法次之,简单随机化最大。结果表明,动态均衡随机化在基线非处理因素的均衡性方面最好,最小化法次之,简单随机化的均衡性最差。

2.2 不平衡检验的显著性 I_{pi} 每个样本量(20、100、200、500、1 000)均重复1 000次随机,测试3种随机化方法导致的组间基线非处理因素不平衡显著性差异的次数,在样本量较小时(20),3种随机化方法都存在不平衡显著性差异,但是简单随机化差异的次数较大。另外,当样本量增大到100及以上时,动态均衡随机和最小化法均不再出现组间非处理因素不平衡显著性差异,而简单随机化的组间非处理因素不平衡显著性差异的次数并没有因为样本量的增加而减少。结果见表4。

2.3 基线症状的组间均衡性 随机选取40例受试者分别进行动态均衡随机、最小化法和简单随机,并比较组间基线症状均衡性,发现除了简单随机在个

表2 各样本量条件下3种随机化方法的组间例数绝对差值 N
Table 2 The absolute difference in number of cases (N) between groups of the three randomization methods under different sample sizes

样本量	均数	25%	50%	75%	极大值
动态均衡随机					
20	0.34	0.00	0.00	0.00	2
100	0.69	0.00	0.00	2.00	4
200	0.55	0.00	0.00	2.00	4
500	0.69	0.00	0.00	2.00	4
1 000	0.88	0.00	0.00	2.00	4
最小化法					
20	0.97	0.00	0.00	2.00	4
100	0.82	0.00	0.00	2.00	4
200	1.56	0.00	2.00	2.00	6
500	1.56	0.00	2.00	2.00	6
1 000	1.37	0.00	2.00	2.00	6
简单随机					
20	3.55	2.00	4.00	6.00	14
100	7.41	2.00	6.00	10.00	30
200	10.89	4.00	10.00	16.00	42
500	14.27	6.00	12.00	20.00	68
1 000	12.67	6.00	10.00	18.00	56

表 3 各样本量条件下 3 种随机化方法的边际不均衡性($\bar{B}_M$, $MaxB_M$)

Table 3 Comparisons of average and maximum marginal imbalance by allocation strategies under different sample sizes

样本量	$\bar{B}_M$					$MaxB_M$				
	均数	25%	50%	75%	极大值	均数	25%	50%	75%	极大值
动态均衡随机										
20	0.199	0.155	0.198	0.243	0.420	0.680	0.333	0.600	1.000	1.000
100	0.043	0.034	0.041	0.051	0.101	0.145	0.091	0.111	0.172	0.667
200	0.021	0.016	0.021	0.025	0.051	0.069	0.043	0.056	0.091	0.294
500	0.008	0.007	0.008	0.010	0.021	0.027	0.018	0.021	0.036	0.123
1 000	0.004	0.003	0.004	0.005	0.009	0.014	0.009	0.010	0.019	0.048
最小化										
20	0.223	0.181	0.222	0.266	0.420	0.740	0.500	0.714	1.000	1.000
100	0.051	0.039	0.050	0.061	0.133	0.188	0.111	0.154	0.238	1.000
200	0.029	0.022	0.027	0.034	0.059	0.104	0.059	0.085	0.125	0.429
500	0.011	0.009	0.011	0.013	0.025	0.042	0.024	0.036	0.055	0.176
1 000	0.006	0.005	0.006	0.007	0.012	0.020	0.013	0.018	0.027	0.074
简单随机										
20	0.330	0.259	0.320	0.394	0.727	0.866	0.667	1.000	1.000	1.000
100	0.142	0.112	0.137	0.166	0.307	0.359	0.273	0.333	0.429	1.000
200	0.100	0.079	0.098	0.119	0.230	0.253	0.183	0.236	0.304	0.733
500	0.061	0.050	0.060	0.071	0.138	0.151	0.113	0.143	0.176	0.391
1 000	0.043	0.037	0.042	0.048	0.069	0.099	0.077	0.093	0.113	0.236

表 4 3 种随机化方法导致的基线组间非处理因素不平衡显著性差异的次数比较

Table 4 Numbers of significant P - values comparing the baseline covariates between groups

随机化方法	不同样本量下 $P<0.05$ 的次数/次				
	20	100	200	500	1 000
动态均衡随机	18	0	0	0	0
最小化法	21	0	0	0	0
简单随机	208	216	267	340	506

别症状组间比较的差异有统计学意义外(见表 5), 其他随机化法基线组间症状比较差异均无统计学意义。而当样本量扩大到 149(广东省中医院芳村分院的例数)时, 所有随机化方法均能达到组间基线症状差异无统计学意义。

2.4 随机性比较 动态均衡随机化和最小化法 2 种方法的第 1 例受试者均采用完全随机化分配入组, 当动态均衡随机化满足 A、B 组中的人数差不超过限值或者最小化法各处理组间总体不均衡性相等时, 也将采用完全随机化分配。这不仅最大程度保证了组间均衡, 而且保证了一定的随机性, 减少试验过程的偏倚^[14]。通过合计最小化法和动态均衡随机化 2 种模拟随机后各自的完全随机次数, 在各样本条件下, 动态均衡随机化的完全随机次数均多于最小化法, 具体见表 6。

3 讨论

本研究结果表明, 在多种样本量(20、100、

表 5 样本量为 40 时简单随机化分配的基线症状组间差异比较

Table 5 Comparison of symptoms between groups at baseline by simple randomization when sample size was 40

指标	分组	例数/例	$\bar{x} \pm s$	t 值	P 值
鼻塞	A	16	1.75 ± 0.93	1.750	0.088
	B	24	1.21 ± 0.98		
咽喉痛	A	16	1.63 ± 0.72	2.675	0.011
	B	24	0.92 ± 0.88		
咳嗽	A	16	1.25 ± 0.68	2.114	0.041
	B	24	0.71 ± 0.86		
肌肉酸痛	A	16	0.75 ± 0.77	-1.363	0.181
	B	24	1.13 ± 0.9		
疲劳	A	16	1.31 ± 0.79	-0.602	0.551
	B	24	1.46 ± 0.72		
头痛	A	16	0.94 ± 0.93	0.486	0.630
	B	24	0.79 ± 0.93		

200、500 和 1 000)条件下, 与最小化法和简单随机化相比, 动态均衡随机化拥有更好的分配均衡性, 具体表现在组间分配例数绝对差值、平均边际不均衡值和最大边际不平衡值。并且, 动态均衡随机化拥有比最小化法更好的随机性。同时, 还发现最小化法的分配均衡性又比简单随机化好。

尽管简单随机化操作相对容易和方便, 也有更好随机性, 但当样本量为 20 时, 用该方法分配有 20% 的概率造成 2 组分配比例为 12:8, 当样本量为 100 时, 该方法造成 2 组比例为 60:40 或更大不均衡的机会略大于 5%^[9, 11]。有学者认为, 在临床试验中选择随机分组方法时, 应该考虑组间均衡性

表 6 各样本量条件下动态均衡随机化和最小化法的完全随机次数

Table 6 number of complete randomization by dynamic balancing randomization and minimization method under different sample sizes

样本量	均数	25%	50%	75%	极大值
动态均衡随机					
20	9.43	8.00	9.00	11.00	15
100	34.08	29.00	34.00	39.00	55
200	63.37	56.00	63.00	71.00	93
500	153.55	140.00	153.50	167.00	215
1 000	293.43	272.00	293.00	314.00	388
最小化法					
20	2.32	2.00	2.00	3.00	6
100	7.79	6.00	8.00	9.00	16
200	14.13	12.00	14.00	16.00	27
500	33.80	30.00	34.00	37.00	53
1 000	65.86	60.00	66.00	71.00	93

(Balance)、预测性(Predictability)和效能(Power)3个因素^[9],而组间均衡性又会影晌统计效能,因而良好的组间均衡性是随机化方法的基本要求^[7]。均衡性不仅包括组间例数的均衡,也包括重要非处理因素的组间分布均衡,而后的严重不均衡给试验造成的偏离可能比前者更严重^[15]。

近年来,随着计算机技术和网络技术的不断发展,动态随机化方法(Dynamic randomization)在临床研究中的应用也越来越多。最小化法在样本量小且存在多个非处理因素时,其组间均衡性远远优于简单随机和分层区组随机^[7, 19]。而本研究通过对3种随机方法的比较,发现动态均衡随机与最小化法相比拥有更好的组间均衡性和随机性,这个结果与相关研究^[8]的结果是一致的。究其原因,应该与动态均衡随机不仅考虑了非处理因素,同时对非处理因素进行了重要程度分级,并设置了各非处理因素允许的组间例数最大差值有关。

综上所述,动态均衡随机和最小化法2种动态随机化方法具有良好的组间均衡性,在小样本、非处理因素较多时比较适用,尤其是当非处理因素之间存在重要程度分级时可以考虑采用动态均衡随机化方法。

参考文献:

[1] FISHER R A. The design of experiments[M]. Edinburgh: Liver and

Boyd, 1935.

- [2] ALTMAN D G, BLAND J M. Treatment allocation in controlled trials: why randomise? [J]. BMJ, 1999, 318(7192): 1209.
- [3] BRIDGMANS, DAINITY K, KIRKLEY A, et al. Practical aspects of randomization and blinding in randomized clinical trials [J]. The Journal of Arthroscopic and Related Surgery, 2003, 19(9): 1000-1006.
- [4] Torgerson D J, Roberts C. Randomisation methods: concealment [J]. BMJ, 1999, 319(7206): 375-376.
- [5] ICH. Statistical principles for clinical trials, E9 [EB/OL]. (1998-07-20) [2019-12-31]. http://www.ich.org/fileadmin/Public_Web_Site/ICH_Products/Guidelines/Efficacy/E9/Step4/E9_Guideline.pdf
- [6] 胡良平, 贾元杰. 用SAS软件实现随机区组设计定量资料的统计分析[J]. 药学服务与研究, 2011, 11(5): 330-332.
- [7] 闫世艳, 姚晨, 夏结来. 简单随机化、中心分层区组随机化和最小化法的均衡性比较[J]. 中国循证医学杂志, 2006, 6(5): 376-379.
- [8] 吴春芳, 许金芳, 陆健, 等. 临床试验中动态均衡随机化方法的均衡性和随机性评价[J]. 中国新药杂志, 2011, 20(7): 590-593.
- [9] TREASURE T, MACRAE K D. Minimization: the platinum standard for trials[J]. BMJ, 1998, 317(8): 362-363.
- [10] POCOCK S J. Allocation of patients to treatment in clinical trials[J]. Biometrics, 1979, 35: 183-197.
- [11] 赖世隆. 中西医结合临床科研方法学[M]. 北京: 科学出版社, 2008: 31-36
- [12] PANDIS N, POLYCHRONOPOULOU A, ELIADES T. Randomization in clinical trials in orthodontics: its significance in research design and methods to achieve it [J]. European Journal of Orthodontics, 2011, 33: 684-690.
- [13] SIGNORINI D F, LEUNG O, SIMES R J, et al. Dynamic balanced randomization for clinical trials[J]. Stat Med. 1993, 12(24): 2343-2350.
- [14] TAVES D R. Minimization: a new method of assigning patients to treatment and control group[J]. Clin Pharmacol Ther, 1974, 15(5): 443-453.
- [15] POCOCK S J, SIMON R. Sequential treatment assignment with balancing for prognostic factors in the controlled clinical trial[J]. Biometrics, 1975, 31(1): 103-115.
- [16] 陆健. 临床试验电子化数据管理与统计分析系统的开发及应用[D]. 上海: 第二军医大学, 2012.
- [17] SCOTT N W, MCPHERSON G C, RAMSAY C R, et al. The method of minimization for allocation to clinical trials: a review[J]. Control Clin Trials, 2002, 23(6): 662-674.
- [18] XIAO L, LAVORI P W, WILSON S R, et al. Comparison of dynamic block randomization and minimization in randomized trials: a simulation study[J]. Clinical Trials, 2011, 8(1): 59-69.
- [19] MCENTEGART D J. The pursuit of balance using stratified and dynamic randomization techniques: an overview[J]. Drug Inf J, 2003, 37: 293-308.

(编辑: 邹元平)