

中药材提取物的混批勾兑研究

刘永锁¹, 曹敏², 陈玉英², 胡育筑², 王义明¹, 罗国安¹

(1. 清华大学化学系 北京 100084; 2. 中国药科大学分析化学教研室 江苏 南京 210009)

摘要 :采用非线性最小二乘拟合计算中药材提取物的勾兑系数,不同批的中药材提取物经过勾兑后与参照样品的差异减小,各成分含量稳定。采用数据预处理的方法,并对数据预处理方法进行改进,使峰面积较小的色谱峰可以实现较小的相对差异。引入误差控制系数,可实现对特定色谱峰的控制要求。实验结果表明,非线性最小二乘拟合可以用于计算中药材提取物的勾兑系数。

关键词 :中药材;勾兑;非线性最小二乘拟合;预处理;误差控制系数

中图分类号 :O658 文献标识码 :A 文章编号 :1000-8713(2006)02-0117-05 栏目类别 :研究论文

Study of Blending Method for the Extracts of Herbal Plants

LIU Yongsuo¹, CAO Min², CHEN Yuying², HU Yuzhu², WANG Yiming¹, LUO Guoan¹

(1. Department of Chemistry, Tsinghua University, Beijing 100084, China;

2. Department of Analytical Chemistry, China Pharmaceutical University, Nanjing 210009, China)

Abstract : The irregularity in herbal plant composition is influenced by multiple factors. As for quality control of traditional Chinese medicine, the most critical challenge is to ensure the dosage content uniformity. This content uniformity can be improved by blending different batches of the extracts of herbal plants. Nonlinear least-squares regression was used to calculate the blending coefficient, which means no great absolute differences allowed for all ingredients. For traditional Chinese medicines, even relatively smaller differences could present to be very important for all the ingredients. The auto-scaling pretreatment was used prior to the calculation of the blending coefficients. The pretreatment buffered the characteristics of individual data for the ingredients in different batches, so an improved auto-scaling pretreatment method was proposed. With the improved auto-scaling pretreatment, the relative differences decreased after blending different batches of extracts of herbal plants according to the reference samples. And the content uniformity control of the specific ingredients could be achieved by the error control coefficient. In the studies for the extracts of fructus gardeniae, the relative differences of all the ingredients is less than 3% after blending different batches of the extracts. The results showed that nonlinear least-squares regression can be used to calculate the blending coefficient of the herbal plant extracts.

Key words : herbal plants; blending; nonlinear least-squares regression; pretreatment; error control coefficient

在中药的质量控制中,质量不稳定是一个非常突出的问题,这个问题产生的主要原因是原药材的质量不稳定,不同产地的原药材中各成分的含量差异较大,同一产地的不同批次之间也存在差异。为了解决这一问题,国外草药制剂生产中采用提取物,特别是“标准提取物”作为产品的“处方原料”,使产品质量趋于稳定,此外,德国 Schwabe 药厂对银杏

叶提取物采用“混批勾兑”的做法,也可以使各种可测成分保持稳定,加上标准的提取工艺,各种可测成分的批间差异不超过 $\pm 5\%$ [1-3]。本文主要对中药材提取物混批勾兑的方法进行研究。

1 原理

中药材提取物混批勾兑首先要考虑需要控制哪

些成分的含量以及各成分含量的控制标准。经中药指纹图谱研究可以得到待勾兑的中药材提取物中各成分的含量(或峰面积)。设需要控制 m 个成分的含量,每个成分的含量以 a 来表示,则 n 个待勾兑样品的 m 个成分的含量可以表示为 $S_1 = (a_{11} a_{12} \dots a_{1m})$, $S_2 = (a_{21} a_{22} \dots a_{2m})$, ..., $S_n = (a_{n1} a_{n2} \dots a_{nm})$ 。设参照样品中 m 个成分的含量控制的标准为 $O = (a_1, a_2, \dots, a_m)$, n 个样品的勾兑系数为 x_1, x_2, \dots, x_n , 则有

$$S_1 \times x_1 + S_2 \times x_2 + \dots + S_n \times x_n = O \quad (x_i \geq 0) \tag{1}$$

中药材提取物勾兑方法的核心就是如何计算勾兑系数,在酿酒工业的勾兑技术中多采用线性规划计算勾兑系数^[4,5],本文采用非线性最小二乘拟合计算勾兑系数,以 Matlab5.3 编程计算,以达到整体误差最小。

2 实验部分

2.1 仪器与试剂

岛津 LC-10AT VP 高效液相色谱仪,SPD-10A VP 检测器,CLASS-VP 色谱工作站,H66025T 型超声清洗器(无锡超声电子设备厂),甲醇(色谱纯,美

国 Tedia 公司),重蒸水(自制)。

2.2 色谱条件

色谱柱:Alltech C_{18} 柱(250 mm \times 4.6 mm i. d. $5 \mu\text{m}$)。流动相 A:水;流动相 B:甲醇;梯度洗脱:初始时 B 的体积分数为 50%,在 50 min 内线性增长至 80%,保持 10 min;流速:1.0 mL/min。检测波长 440 nm。进样量 20 μL 。

2.3 药材

江西道地产区栀子药材 10 批。

2.4 样品制备

将栀子药材粉碎,过 20 目筛,取适量(约 1 g),精密称定,加 75% 乙醇 10 mL,超声处理 1 h,过滤,残渣加 75% 乙醇 10 mL 再超声处理 1 h,过滤,合并滤液。吸取滤液 5 mL 置于 25 mL 量瓶中,加水稀释至刻度,用 0.45 μm 微孔滤膜过滤,滤液作为供试品溶液。

3 结果与讨论

经高效液相色谱分析,10 批栀子药材的提取物有 7 个色谱峰。表 1 中是 10 批栀子提取物的 7 个色谱峰的峰面积和参照样品的 7 个色谱峰的峰面积(10 个样品的中位数)及其峰面积的相对标准偏差。

表 1 10 批栀子药材提取物和参照样品的 7 个色谱峰的峰面积($\times 10^4$)及其相对标准偏差(RSD)
Table 1 Peak areas ($\times 10^4$) of 10 batches of the extracts of fructus gardeniae and the reference sample and their relative standard deviations (RSD)

Peak No.	Sample No.										Reference sample	RSD/%
	1	2	3	4	5	6	7	8	9	10		
1	48832	39233	41840	26942	28479	26970	40475	31856	49286	76740	39854	34.9
2	1754872	1595073	1696835	1103863	1407074	1439469	1503337	1364387	1889388	1283671	1471403	15.0
3	291716	248303	256658	141663	160330	137030	238236	196581	304385	247842	243039	25.6
4	52058	42285	44765	36547	44965	37190	41510	31331	64180	62091	43525	22.5
5	174827	83616	103233	122169	142688	129422	82727	71585	133546	90812	112701	27.3
6	25965	19113	19664	19372	22696	17609	27819	13671	29703	48973	21180	39.2
7	29629	28219	25909	37989	44352	38648	27186	15831	28173	93869	28924	56.7

Peak areas of reference sample are the medians of peak areas of 10 batches of the extracts of fructus gardeniae.

3.1 栀子提取物与参照样品的差异

以参照样品中各色谱峰的峰面积作为质量控制标准,10 批栀子提取物能否直接用于制剂的生产,要看它们与参照样品的相似程度或差异大小。为了分析 10 批栀子提取物与参照样品的相似程度,采用相关系数和夹角余弦^[6]计算 10 个样品与参照样品的相似度。以相关系数计算的 10 个样品的相似度依次为 0.999 7, 0.999 7, 0.999 9, 0.998 3, 0.997 8, 0.997 1, 0.999 7, 0.999 7, 0.999 9, 0.998 8;以夹角余弦计算的 10 个样品的相似度依次为 0.999 7, 0.999 7, 0.999 8, 0.998 7, 0.998 3, 0.997 5, 0.999 7, 0.999 4, 1.000 0, 0.997 3。

由相关系数和夹角余弦的计算结果可以看出,10 批样品与参照样品的相似度很高,全部在 0.99 以上,似乎不需要勾兑就可以直接应用。但比较 10 批样品和参照样品的峰面积会发现存在较大的差异。为了便于比较差异的大小,我们计算了 10 批样品与参照样品的各峰面积的百分比,结果见表 2。由计算结果可以看出,10 批栀子提取物的峰面积与参照样品的峰面积存在较大的相对差异,而相关系数和夹角余弦并不能客观地全面反映样品的差异。对于中药而言,要保证药效稳定,各成分的含量相对差异不能太大,因此需要勾兑不同批中药的提取物才能减小与参照样品的差异。

表 2 10 批栀子提取物与参照样品峰面积的百分比

Table 2 Ratios of the peak areas of 10 batches of the extracts of fructus gardeniae to the reference sample %

Peak No.	Sample No.									
	1	2	3	4	5	6	7	8	9	10
1	122.5	98.44	105.0	67.60	71.46	67.67	101.6	79.93	123.7	192.6
2	119.3	108.4	115.3	75.02	95.63	97.83	102.2	92.73	128.4	87.24
3	120.0	102.2	105.6	58.29	65.97	56.38	98.02	80.88	125.2	102.0
4	119.6	97.15	102.9	83.97	103.3	85.45	95.37	71.98	147.5	142.7
5	155.1	74.19	91.60	108.4	126.6	114.8	73.40	63.52	118.5	80.58
6	122.6	90.24	92.84	91.46	107.2	83.14	131.4	64.55	140.2	231.2
7	102.4	97.56	89.58	131.3	153.3	133.6	93.99	54.73	97.40	324.5

3.2 勾兑后的样品

以非线性最小二乘拟合计算 10 批栀子提取物的勾兑系数,目标为参照样品,结果依次为 0.347 6, 0.199 6, 0.140 8, 0.000 0, 0.000 0, 0.000 0, 0.000 5, 0.000 0, 0.119 5, 0.060 9。

按照上述勾兑系数混合 10 批栀子提取物可以得到与参照样品整体差异最小的混合样品,混合样品各色谱峰的峰面积以及它们与参照样品峰面积的绝对差异和相对差异见表 3。

表 3 混合样品的各色谱峰的峰面积以及与参照样品的差异

Table 3 Peak areas of the blend sample and differences according to the reference

Peak No.	Peak area of blend sample/10 ⁴	Absolute difference/10 ⁴	Relative difference/%
1	41277	1423	3.57
2	1471974	571	0.04
3	238679	-4360	-1.79
4	44308	783	1.80
5	113522	821	0.73
6	22153	973	4.59
7	28673	-251	-0.87

由表 3 的计算结果可以看出,勾兑后的混合样品与参照样品的各色谱峰的峰面积的绝对差异和相对差异与勾兑前相比显著减小,各色谱峰的峰面积

表 4 10 批栀子提取物和参照样品的峰面积范围标度化^[6]预处理的结果

Table 4 Results of 10 batches of the extracts of fructus gardeniae and the reference sample after auto-scaling^[6] pretreatment

Peak No.	Sample No.										Reference sample	RSD/%
	1	2	3	4	5	6	7	8	9	10		
1	0.4396	0.2468	0.2992	0.0000	0.0309	0.0006	0.2718	0.0987	0.4487	1.0000	0.2593	102.2
2	0.8288	0.6253	0.7549	0.0000	0.3860	0.4272	0.5085	0.3317	1.0000	0.2289	0.4679	56.3
3	0.9243	0.6649	0.7148	0.0277	0.1392	0.0000	0.6047	0.3558	1.0000	0.6621	0.6334	65.9
4	0.6310	0.3335	0.4090	0.1588	0.4151	0.1784	0.3099	0.0000	1.0000	0.9364	0.3712	72.3
5	1.0000	0.1165	0.3065	0.4900	0.6887	0.5602	0.1079	0.0000	0.6002	0.1862	0.3982	74.0
6	0.3483	0.1542	0.1698	0.1615	0.2557	0.1116	0.4008	0.0000	0.4541	1.0000	0.2127	90.2
7	0.1768	0.1587	0.1291	0.2839	0.3655	0.2924	0.1455	0.0000	0.1582	1.0000	0.1678	100.8

由表 4 中的数据可以看出,采用范围标度化预处理,若样品的色谱峰的峰面积为最小值时,预处理后它的峰面积会变为零,而峰面积为零时,不论混合

与参照样品相比,相对差异不超过 ± 5%。

与文献[4,5]中的线性规划相比,采用非线性最小二乘拟合计算中药材提取物的勾兑系数,所选择的各个成分都参与质量控制,不会因为某些成分的含量不合格而舍弃对它们的质量控制要求。

3.3 数据的预处理

以非线性最小二乘拟合计算勾兑系数,要求离差平方和最小,即不允许出现较大的绝对差异。相同的绝对差异对于峰面积较大的峰可能影响不大,但对于峰面积较小的峰影响较大,不能满足质量控制要求。为了解决这个问题可以采取两种做法,一是加权最小二乘法,即允许峰面积较大的峰有较大的绝对差异,而峰面积较小的峰有较小的绝对差异;二是进行数据预处理,即放大小峰的差异,减小大峰的差异。本文采用数据预处理的办法。在数据的预处理方法中,范围标度化^[6]预处理能达到上述目的,它采用公式(2)进行计算

$$x_{ij, new} = (x_{ij, old} - x_{ij, 最小}) / (x_{ij, 最大} - x_{ij, 最小}) \quad (2)$$

其中 $x_{ij, new}$ 为数据预处理后新生成的变量, $x_{ij, old}$ 为数据预处理前的原始变量, $x_{ij, 最大}$ 和 $x_{ij, 最小}$ 为数据预处理前相同变量中的最大和最小分量。

按公式(2)处理表 1 中的数据,结果见表 4。

系数为多大,对混合样品没有任何影响。而实际情况并非如此,经过范围标度化预处理后,与原始数据相比,10 批栀子提取物和参照样品的相同色谱峰的

RSD 发生了变化,即相同色谱峰的峰面积的分布特征改变了。

为了克服这一缺点,对范围标度化预处理公式进行了改进,改进后的预处理公式如下:

$$x_{ij,new} = x_{ij,old} / (x_{ij,最大} - x_{ij,最小}) \quad (3)$$

表 5 以改进的预处理方法对 10 批栀子提取物和参照样品峰面积的处理结果

Table 5 Results of 10 batches of the extracts of fructus gardeniae and the reference sample after improved auto-scaling pretreatment

Peak No.	Sample No.										Reference sample	RSD/%
	1	2	3	4	5	6	7	8	9	10		
1	0.9806	0.7878	0.8402	0.5410	0.5719	0.5416	0.8128	0.6397	0.9897	1.5410	0.8003	34.9
2	2.2340	2.0306	2.1601	1.4053	1.7913	1.8325	1.9138	1.7369	2.4053	1.6342	1.8731	15.0
3	1.7431	1.4837	1.5336	0.8465	0.9580	0.8188	1.4235	1.1746	1.8188	1.4809	1.4522	25.6
4	1.5848	1.2873	1.3628	1.1126	1.3688	1.1322	1.2637	0.9538	1.9538	1.8902	1.3250	22.5
5	1.6934	0.8099	0.9999	1.1833	1.3821	1.2536	0.8013	0.6934	1.2935	0.8796	1.0916	27.3
6	0.7355	0.5414	0.5570	0.5488	0.6429	0.4988	0.7880	0.3873	0.8414	1.3873	0.6000	39.2
7	0.3797	0.3616	0.3320	0.4868	0.5683	0.4952	0.3484	0.2029	0.3610	1.2029	0.3706	56.7

3.4 改进预处理方法后计算的勾兑系数

以表 5 中的数据采用非线性最小二乘拟合计算 10 批栀子提取物的勾兑系数,结果依次为 0.335 7, 0.279 2, 0.149 9, 0.027 0, 0.002 6, 0.000 0, 0.000 5, 0.000 0, 0.058 4, 0.041 0。

根据此勾兑系数混合 10 批栀子提取物得到混合样品,各色谱峰的峰面积及其与参照样品峰面积的绝对差异和相对差异见表 6。

表 6 以改进的预处理方法处理后所得的混合样品的各色谱峰的峰面积及其与参照样品的差异

Table 6 Peak areas of the blend sample and differences according to the reference after improved auto-scaling pretreatment

Peak No.	Peak area of blend sample/10 ⁴	Absolute difference/10 ⁴	Relative difference/%
1	40464	610	1.53
2	1485943	14540	0.99
3	238017	-5022	-2.07
4	43409	-116	-0.27
5	112742	41	0.04
6	21338	158	0.75
7	28358	-566	-1.96

表 7 以改进的预处理方法处理 4~10 号样品和参照样品的峰面积的结果

Table 7 Results of batches 4-10 of the extracts of fructus gardeniae and the reference sample after improved auto-scaling pretreatment

Peak No.	Sample No.							Reference sample
	4	5	6	7	8	9	10	
1	0.5410	0.5719	0.5416	0.8128	0.6397	0.9897	1.5410	0.8003
2	1.4053	1.7913	1.8325	1.9138	1.7369	2.4053	1.6342	1.8731
3	0.8465	0.9580	0.8188	1.4235	1.1746	1.8188	1.4809	1.4522
4	1.1126	1.3688	1.1322	1.2637	0.9538	1.9538	1.8902	1.3250
5	1.7182	2.0068	1.8202	1.1635	1.0068	1.8782	1.2772	1.5850
6	0.5488	0.6429	0.4988	0.7880	0.3873	0.8414	1.3873	0.6000
7	0.4868	0.5683	0.4952	0.3484	0.2029	0.3610	1.2029	0.3706

按式(3)对表 1 中的数据进行预处理,结果见表 5。

以改进后的预处理方法对 10 批栀子提取物和参照样品的峰面积处理后,没有零值出现,而且与原始数据相比,相同色谱峰的 RSD 不变,即相同色谱峰的峰面积分布特征没有改变。

由表 6 中的数据可以看出,经改进的数据预处理方法处理之后,峰面积较小的色谱峰的绝对差异减小,峰面积较大的峰的绝对差异增大,但整体的相对差异控制在 ±3% 以内,比预处理前效果改善。

通过数据预处理,可以实现对各成分的相对含量的控制,而不是以绝对差异控制各成分的含量,这符合药物的量效关系。

3.5 剩余样品的处理

由“3.4”节计算出的勾兑系数可以看出,有几个样品在勾兑时用量比较多,有些则用量比较少,而有些样品几乎不用。下面去除批号为 1, 2, 3 的 3 个样品,以其他的几个样品勾兑。首先以改进的预处理方法处理 4~10 号样品和参照样品的峰面积,结果见表 7。

以表 7 中的数据计算 4~10 号样品的勾兑系数,结果依次为 0.296 7, 0.000 0, 0.000 0, 0.065 7, 0.361 3, 0.312 8, 0.008 6。按此勾兑系数混合样品,所得混合样品的峰面积以及它与参照样品峰面积的绝对差异和相对差异见表 8。

表 8 改进预处理方法后所得混合样品的峰面积
及其与参照样品的差异

Table 8 Peak areas of the blend sample and differences according to the reference sample after improved auto-scaling pretreatment

Peak No.	Peak area of blend sample/ 10^4	Absolute difference/ 10^4	Relative difference/%
1	38231	-1623	-4.07
2	1520996	49592	3.37
3	226006	-17033	-7.01
4	45491	1966	4.52
5	110081	-2620	-2.32
6	22221	1041	4.92
7	28390	-534	-1.85

由上述计算结果可以看出,用 4~10 号样品勾兑,与用 1~10 号样品相比,所得到的混合样品与参照样品的差异变大,最大的差异超过 7%,但比没有勾兑之前差异仍然减小。

因此以非线性最小二乘拟合做中药的勾兑,得到的是与参照样品整体差异最小的混合样品,而不能控制整体差异一定在某一范围之内。混合样品与参照样品差异的大小由样品本身的差异决定。

将“3.4”节 1~10 号栀子提取物混合得到的样品和“3.5”节 4~10 号栀子提取物混合得到的样品以 1:1.4 比例混合时,可以得到与参照样品最大相对差异不超过 $\pm 5\%$ 的样品。

3.6 引入误差控制系数

对于中药的质量控制,由于不同成分的药理作用不同,对某些成分的含量控制可能比较严格,而对另外一些成分的含量控制可能比较宽松。如果经过勾兑后某个或某几个色谱峰与参照样品的差异仍然比较大,不符合质量控制要求,则为了实现特定色谱峰的控制要求,可以引入误差控制系数。

引入误差控制系数的方法可以参照数据的预处理方法,数据经过预处理后,采用非线性最小二乘拟合计算勾兑系数得到的混合样品中差异被放大的色谱峰与参照样品中相应的色谱峰有较小的绝对偏差,而差异被缩小的色谱峰与参照样品中相应的色谱峰有较大的绝对偏差。因此如果某个色谱峰与参照样品中相应色谱峰的差异较大,在预处理时可以把差异进一步放大。在计算时可以采用公式(4)。

$$x_{ij, \text{new}} = f \cdot x_{ij, \text{old}} / (x_{ij, \text{最大}} - x_{ij, \text{最小}}) \quad (4)$$

误差控制系数 f 直接控制色谱峰差异改变的情况,当差异放大以后,混合样品中的色谱峰与参照样品中相应的色谱峰的差异会减小,但同时其他峰的差异会增大。

引入误差控制系数可以通过调整其他成分的含量来实现对特定色谱峰的含量控制要求,避免某些成分含量不合格时舍弃对它们的控制要求。

3.7 存在的问题

以非线性最小二乘拟合计算中药的勾兑系数,理论上可以得到与标准样品或参照样品差异最小的混合样品,但不同的优化算法只能找到近似的最小值而不是理论上的最小值。非线性最小二乘拟合结合不同的优化算法均存在局部最小值的问题,我们曾结合遗传算法,但计算的误差较大,而结合高斯-牛顿算法进行迭代运算,可以得到较小的误差。另外,以最小二乘原则计算勾兑系数所得到的混合样品与参照样品的差异最小,对于中药而言只要把各成分的含量变化控制在一定的范围内即可,比如 10% 的变化,或者 20% 的变化,也可能有些可以大些,有些要小些,因此中药的混批勾兑方法应该依照这种思路继续完善。

4 结论

按照非线性最小二乘拟合计算的中药材提取物的勾兑系数混合不同批次的样品,可以减小批间差异,使混合后的样品与参照样品的绝对差异和相对差异都减小;以改进的范围标度化法对数据进行预处理,相同变量的预处理后的数据和原始数据的分布特征不变,而且可以使峰面积较小的色谱峰实现较小相对误差;引入误差控制系数可以实现对特定色谱峰的峰面积控制要求。将该法用于制剂的生产,可以提高中药质量的稳定性。

参考文献:

- [1] Xie Peishan. World Science and Technology/Modernization of Traditional Chinese Medicine (谢培山. 世界科学与技术/中药现代化), 2001, 3(3): 18
- [2] Luo Guoan, Wang Yiming. Chinese Journal of New Drugs (罗国安, 王义明. 中国新药杂志), 2002, 11(1): 46
- [3] Jin Zhangzhao, Zhu Ming, Wu Wenjun, Yu Jianping, Ma Linke. Chinese Traditional Patent Medicine (金樟照, 祝明, 吴文军, 俞剑平, 马临科. 中成药), 2003, 25(10): 777
- [4] Huang Hanying, Wei Mingxin, Yao Jicheng. Journal of Huazhong Agriculture University (黄汉英, 魏明新, 姚继承. 华中农业大学学报), 2001, 20(4): 400
- [5] Zhu Jinyan, Hou Jingde, Feng Jianyue. Analysis and Testing Technology and Instruments (朱金炎, 侯镜德, 冯建跃. 分析测试技术与仪器), 1997, 3(3): 135
- [6] Liang Yizeng, Yu Ruqin. Handbook of Analytical Chemistry (2nd ed), Vol. 10: Chemometrics. Beijing: Chemistry Industry Press (梁逸曾, 俞汝勤. 分析化学手册(第2版), 第十分册: 化学计量学. 北京: 化学工业出版社), 2000. 328