

统计过程控制在栀子前处理生产工艺中的应用

周海燕^{1,3}, 徐冰^{1,2}, 史新元^{1,2*}, 乔延江^{1,2*}
(1. 北京中医药大学中药学院, 北京 100102;
2. 国家中医药管理局中药信息工程重点研究室, 北京 100102;
3. 亚宝北中大(北京)制药有限公司, 北京 101300)

[摘要] 目的:以清开灵注射液中栀子前处理工艺为载体,研究统计过程控制在中药生产过程中的应用。方法:利用 Shewhart 控制图和 EWMA 控制图对栀子苷提取率进行单变量统计过程控制研究;在主成分分析模型的基础上,建立栀子前处理工艺多变量 Hotelling T² 控制图和 SPE 控制图;利用贡献图对工艺过程进行诊断。结果:Shewhart 控制图与 EWMA 控制图可实现栀子前处理工艺中质量参数的监控;多变量统计过程控制图可实现多个过程参数的监控,结合贡献图可实现工艺分析和故障诊断。结论:单变量统计过程控制与多变量统计过程控制的结合,可更好地监控并理解中药生产过程。

[关键词] 统计过程控制;控制图;栀子;生产工艺;主成分分析;工艺诊断

[中图分类号] R283.6 [文献标识码] A [文章编号] 1005-9903(2012)11-0016-05

Application of Statistical Process Control in Pretreatment Production Process of *Gardenia jasminoides*

[收稿日期] 20111202(009)
[基金项目] 国家“重大新药创制”科技重大专项(2010ZX09502-002);北京市支持中央在京高校共建项目;北京中医药大学科研基金项目(JYB22-XS034)
[第一作者] 周海燕,在读博士,从事中药质量控制研究, E-mail: zhouhaiyan891@sohu.com
[通讯作者] * 乔延江,教授,博士生导师,从事中药信息工程研究, Tel:010-84738620, E-mail: yjqiao@263.com; * 史新元,副教授,硕士生导师,从事中药过程分析研究, Tel:010-84738621, E-mail: shixinyuan01@163.com

被破坏,提取率高达 90%,适合于大规模生产,可作为鲜地黄药材特有的加工提取方法。

[参考文献]

[1] 刘彦飞,赵羽,温学森,等. 地黄化学成分及其在加工炮制过程中的变化[J]. 国外医药:植物药分册, 2007,22(7):102.

[2] 赵宇,温学森,武卫红. 地黄不同炮制品中梓醇含量分析现状[J]. 中国药学杂志,2007,42(7):486.

[3] 王宏洁,边宝林,杨健. 地黄中梓醇变化条件的探讨[J]. 中国中药杂志,1997,22(7):408.

[4] 赵宇,温学森,崔晶,等. 鲜地黄中 α -葡萄糖苷酶和 β -葡萄糖苷酶的提取与初步纯化[J]. 中药材,2006,29(2):137.

[5] LI G S, LIU C H, WANG H S, et al. Assaying of content of catalpol in *Rehmannia glutinosa* from different origins [J]. Chin Tradit Herb Drugs, 2002, 33 (2): 126.

[6] LUO Y Y, ZHANG S Q, SUO J Z, et al. Determination of catalpol in Radix Rehmanniae by high performance liquid chromatography [J]. Chin Pharm J, 1994, 29 (1): 38.

[7] Tomoda M, Kato S, Onuma M. Water-soluble constituents of Rehmanniae Radix I carbohydrates and acids of Rehmannia Glutinosafhueioh ingensis [J]. Chem Pharm Bull, 1971, 19 (7): 1455.

[8] 刘延泽. 植物组织破碎提取法及闪式提取器的创制与实践[J]. 中国天然药物, 2007,5(6): 401.

[9] 吴冬梅. 闪式提取器在中药研究中的应用[J]. 中国实验方剂学杂志, 2006,12(7): 34.

[10] 刘明,李更生. 鲜地黄中梓醇提取工艺[J]. 时珍国医国药, 2000, 11(4): 301.

[11] 钱桂敏,王平,郭峰. 金钗石斛鲜品闪式提取方法的工艺研究[J]. 中国实验方剂学杂志, 2008, 14 (5): 38.

[12] 贺石麟,牛景霞,倪艳. 淫羊藿中淫羊藿苷和总黄酮的闪式提取工艺[J]. 中国实验方剂学杂志, 2011, 17(8): 38.

[责任编辑 仝燕]

ZHOU Hai-yan^{1,3}, XU Bing^{1,2}, SHI Xin-yuan^{1,2*}, QIAO Yan-jiang^{1,2*}

- (1. School of Chinese Medicine, Beijing University of Chinese Medicine, Beijing 100102, China;
2. Key Laboratory of Traditional Chinese Medicine Information Engineering, State Administration
of Traditional Chinese Medicine, Beijing 100102, China;
3. Yabao Beizhongda (Beijing) Pharmaceutical Co. Ltd, Beijing 101300, China)

[Abstract] **Objective:** To study on application of statistical process control in pretreatment production process of Chinese medicine with pretreatment production process of *Gardenia jasminoides* from Qingkailing injection as carrier. **Method:** Univariate statistical process control for extraction ratio of geniposide was researched by Shewhart control chart and EWMA control chart; Based on principal component analysis model, multivariate Hotelling T² and SPE control charts for pretreatment process of *G. jasminoides* were built; Technology process was diagnosed by contribution chart. **Result:** Shewhart and EWMA control charts could be achieved monitoring quality parameters in pretreatment technology of *G. jasminoides*; Multivariate statistical process control chart could be successfully applied in monitoring multiple process parameters; process analysis and fault diagnosis could be achieved by combining with contribution chart. **Conclusion:** Combination of univariate and multivariate statistical process control could be well monitored and understood of production process for traditional Chinese medicine.

[Key words] statistical process control; control chart; *Gardenia jasminoides*; production process; principal component analysis; process diagnosis

中药生产线在长期运行过程中,由于受到各种因素的影响,会不可避免地发生各种变化或漂移,从而引起中药制剂产品质量的改变,甚至危害生命。因此,有必要在现有中药生产过程体系中,建立一套合理的过程监控策略,以保证中药产品质量的均一性和稳定性,进而保证其安全有效。

统计过程控制 (SPC) 又称统计质量控制 (SQC),随着全面质量管理 (TQM) 思想^[1]的普及,SPC 在企业中应用逐渐从生产制造过程质量控制扩展到产品设计、辅助生产过程、售后服务及产品使用等各个环节的质量控制,强调全过程的预防与控制^[2]。要保证中药临床安全有效、质量稳定可控,就必须加强中药生产全过程质量控制研究^[3]。中药过程质量控制就是对中药生产过程的关键环节进行及时测量,并对生产过程进行反馈和控制^[4],与 SPC 最终目标是一致的。

本文以清开灵注射液中栀子前处理生产工艺为载体,研究单变量统计过程控制 (univariate SPC, USPC) 和多变量统计过程控制 (multivariate SPC, MSPC) 方法学及其在中药生产过程中的应用,分析并比较 2 种监控策略的作用及特点,以期对统计过程控制在中药领域的应用提供依据。

1 数据来源和预处理

1.1 数据来源 利用实际生产过程的数据建立统计分析模型,是实现统计过程控制的关键。收集了

2008 ~ 2010 年亚宝北中大 (北京) 制药有限公司清开灵注射液中栀子提取液制备过程,共 117 批历史生产数据。单批栀子提取液生产过程包括 8 个工艺单元 (A ~ H)、32 个工艺参数变量和栀子苷提取率 (前处理工艺所得中间体中栀子苷总量与该批栀子药材中栀子苷总量的比值) 变量,每个单元对应的工艺参数以大写英文字母加数字表示 (表 1)。

表 1 清开灵注射液中栀子前处理工艺单元及工艺参数

工艺单元	工艺参数
A	A
B	B ₁ , B ₂
C	C ₁ , C ₂ , C ₃ , C ₄
D	D ₁ , D ₂ , D ₃
E	E ₁ , E ₂ , E ₃ , E ₄ , E ₅ , E ₆ , E ₇
F	F ₁ , F ₂ , F ₃ , F ₄ , F ₅ , F ₆
G	G ₁ , G ₂ , G ₃ , G ₄ , G ₅ , G ₆
H	H ₁ , H ₂ , H ₃

1.2 分析软件 数据处理工作在 Matlab 7.12 (美国 MathWorks 公司) 平台下完成,相关分析程序均为自主编制。

1.3 样本划分和数据预处理 将 117 批历史生产数据整理成 117 行 32 列的二维矩阵 *X*, 栀子苷提取率组成单列矩阵 *Y*。采用 Kennard-Stone 算法进行样本划分,距离样本中心最远的点优先选出。取 3/4 作为训练集 (*X_{train}*, *Y_{train}*, 87 个批次), 1/4 作为测试集 (*X_{pre}*, *Y_{pre}*, 30 个批次)。对所有数据按照公式 (1) 进行均值标准化处理,以消除不同工艺参数之间的量

纲差异。

$$XS = \frac{X - X_{mean}}{S}$$
公式(1)

X_{mean} 为 X 的均值矩阵, S 为 X 的标准差矩阵。

2 单变量统计过程控制

USPC 是对生产过程中的一些重要指标单独地实施统计过程控制, 本文重点研究栀子苷提取率的 USPC 方法。假设栀子苷浓度 y_i 服从正态分布, 当 y_i 符合规定时, 其均值大小以及分布曲线轮廓保持一致。通过控制图即可得出栀子苷提取率, 处于被控状态或失控状态的结论。

2.1 Shewhart 控制图

其控制限计算方法如下:

$$UCL = \mu + k\sigma$$
公式(2)

$$LCL = \mu - k\sigma$$
公式(3)

μ 为 Y_{train} 均值, σ 为 Y_{train} 标准差, k 通常取 2 或 3 (本文中 k 取 3)。因为数据经过均值标准化处理, 因此训练集的均值为 0, 标准差为 1, 控制限为 ± 3 。将 Y_{pre} 按顺序绘制于 Shewhart 控制图上(图 1)。

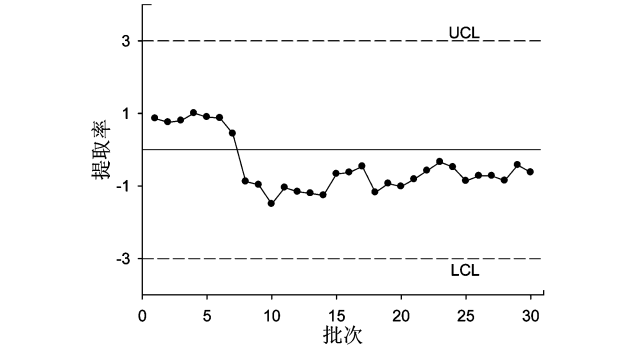


图 1 清开灵注射液中不同批次栀子
提取液制备过程 Shewhart 控制

2.2 EWMA 控制图

指数加权移动平均(EWMA)图适用于过程为“每次一个数据”的情况, 或需要高灵敏度来检测过程微小变化^[5]。其优点是能很好的检测出小的偏移。

$$EWMA_t = \lambda Y_t + (1 - \lambda) EWMA_{t-1} \quad t = 1, 2, \dots, n;$$
公式(4)

$EWMA_0$ 为训练集均值 0, Y_t 为预测集第 t 批的提取率($n=30$), λ 为权重, 其取值范围通常在 0.2 ~ 0.3 之间, 本文中 λ 取 0.3。EWMA 控制限的计算方法如下:

$$UCL = EWMA_0 + ks \sqrt{\frac{\lambda}{2 - \lambda}}$$
公式(5)

$$LCL = EWMA_0 - ks \sqrt{\frac{\lambda}{2 - \lambda}}$$
公式(6)

式中 k 的取值同 Shewhart 控制图中 k 的取值, s 为标准差 1。按照公式计算预测集中各批次提取率

的 EWMA 值, 按照公式计算控制限(± 1.2603), 绘制 EWMA 控制图见图 2。

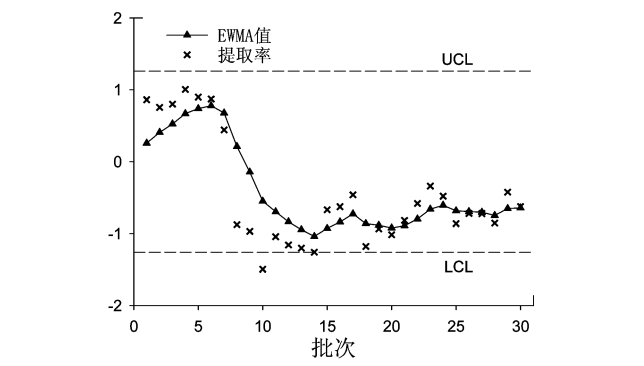


图 2 清开灵注射液中不同批次栀子
提取液制备过程 EWMA 控制

由图 1~2 可知, 训练集中 30 批次栀子苷提取率的变化保持在控制限之内, 即 30 批均为正常合格批次。然而, 传统的单变量统计过程控制技术的局限性在于仅注意监测少量的质量变量及与质量有关的过程变量。对于质量参数的监控, 一旦发现问题并不能找出问题发生的根源所在。本文中栀子前处理工艺涉及 32 个过程变量, 若全部采用 USPC 法, 需要 32 张控制图, 不利于操作人员实际应用。因此关键在于将多变量统计过程控制技术引入中药生产过程体系。

3 多变量统计过程控制

在实际中药生产中, 过程参数(包括工艺参数和质量参数)不只 1 个, 而且相互关联。MSPC 通过多元统计投影(如 PCA, PLS 等)的方法, 可以将中药生产过程中存在的大量高度相关过程变量映射到由少量潜变量定义的低维空间, 从而使过程监控、故障检测和诊断以及相应的研究得以简化^[6]。

3.1 主成分分析

主成分分析统计过程监测模型用以描述正常工况下各过程变量间的关联。建立模型前, 采用主成分贡献率累积和百分比(CPV)方法, 计算各主成分贡献率, 选择累积和百分比 > 85% 的主成分个数。如图 3 所示, 11 个主成分可以解释 X_{train} 中 87.18% 的变异, 因此选择 11 个主成分建立 MSPC 监控模型。

3.2 MSPC 控制限的建立

MSPC 的重要工具是 MSPC 图, 本文主要研究 Hotelling T^2 图和平方预测误差(SPE)图在栀子前处理工艺中的应用。利用主成分模型对生产过程进行监控, 需要由过程正常历史生产的数据(本文中即训练集数据)来确定 Hotelling T^2 和 SPE 统计量的控制限^[7]。

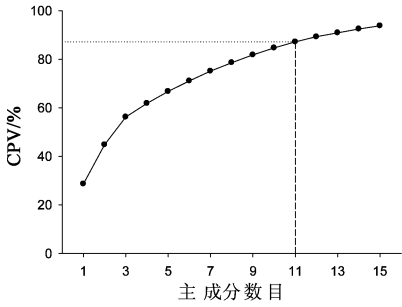


图3 多变量统计过程控制主成分数的选择

Hotelling T^2 (又称 D 统计量, D-statistic) 控制限的计算方法如下:

$$T_{lim}^2 = \frac{K(N-1)}{N-K} F(K, N-K, \alpha) \quad \text{公式(7)}$$

式中 N 是样本个数, K 是主成分模型中所保留的前 K 个主元个数, α 是检验水平, F 为 F 分布临界值。SPE (又称 Q 统计量, Q -statistic) 的计算方法如下:

$$SPE_{\alpha} = \theta_1 \left[\frac{c_{\alpha} h_0}{\theta_1} \sqrt{2\theta_2} + 1 + \frac{\theta_2 h_0 (h_0 - 1)}{\theta_1^2} \right]^{\frac{1}{h_0}} \quad \text{公式(8)}$$

$$\theta_i = \sum_{j=K+1}^m \lambda_j^i h_0 = 1 - \frac{2\theta_1 \theta_3}{3\theta_2^2}$$

式(8)中, λ 为 X_{train} 的协方差矩阵的特征值, C_{α} 是正态分布在检验水平为 α 下的临界值, K 同公式(7)中主元模型中保留的主成分个数, m 是由 X_{train} 分解产生的主成分个数。

取临界水平 $\alpha = 0.95$ 计算警戒限, 由 87 批栀子前处理工艺参数确定的 Hotelling T^2 警戒限为 23.861 7, SPE 警戒限为 7.5814。取临界水平 $\alpha = 0.99$ 计算控制限, 由训练集工艺参数确定的 Hotelling T^2 控制限为 30.998 7, SPE 控制限为 9.822 3。

3.3 MSPC 控制模型的应用 一般认为, Hotelling T^2 描述由统计模型所决定的前 k 个隐变量的综合波动程度, SPE 则描述生产过程和统计模型的偏离程度。参照文献[8], 计算测试集 30 个批次的 Hotelling T^2 统计量和 SPE 统计量, 绘制 MSPC 监控图(图 4)。

Hotelling T^2 监控结果表明, 测试集 30 批工艺参数的 Hotelling T^2 值均在警戒限以内, 说明由主成分解释的过程变量运行正常。SPE 监控图中, 有 5 批数据超过了控制限, 当 SPE 统计量发生较大变化时, 表明 PCA 模型所代表的正常工作状态下的变量

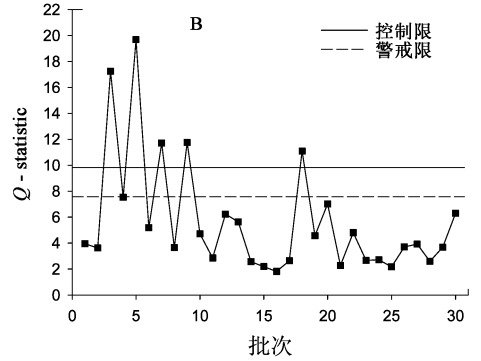
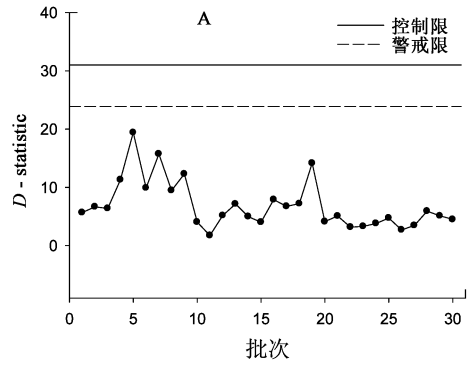
A. Hotelling T^2 控制; B. SPE 控制

图4 30 批次栀子提取液制备过程 MSPC 监控

之间的关系或相关结构被破坏了, 导致这种变化的原因需要进一步工艺诊断。

4 工艺诊断

当 Q 统计量超出其控制限, 可判断过程中出现了不正常情况。但不能从 Q 统计图直接找出故障源, 而贡献图则能帮助确定故障源^[9]。第 j 个过程变量对第 i 个样本的 Q 统计量的贡献如下:

$$Q_{ij} = e^2 = (X_{ij} - \hat{X}_{ij})^2 \quad \text{公式(9)}$$

式中 X_{ij} 为第 i 个样本第 j 个变量的测量值, $\hat{X}_{ij}$ 为第 i 个样本第 j 个变量的主元模型预测值。过程变量的贡献值越大, 说明该变量的变化对于 SPE 统计量的影响越大。结合过程工艺分析贡献值大的过程变量可能对生产过程造成的影响, 即可定位引起过程异常的原因。以批次 5 为例, 各个过程变量对 Q 统计量的贡献直方图见图 5。

由图 5 可知, 导致批次 5 的 Q 统计量超过控制限的最主要的过程参数是 H_2 , 其次是 F_1 和 D_1 。Shewhart 监控结果显示第 10 批栀子苷提取率偏低, 同样可将批次 10 各个过程变量对 Q 统计量的贡献绘制成直方图(图 6)。

由图 6 可知, 导致批次 10 栀子苷提取率偏低的原因主要由过程参数 F_1 引起, 其次为 H_2 , A 和 G_5 。质量控制人员可查看这些过程参数的历史记录, 进

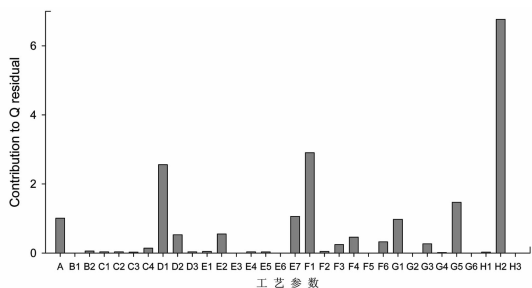


图 5 验证集第 5 批栀子提取液制备过程 SPE 贡献

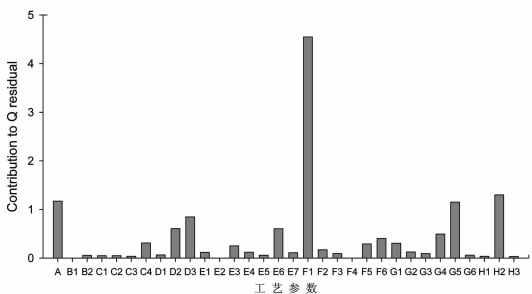


图 6 验证集第 10 批栀子提取液制备过程 SPE 贡献

一步分析是系统故障,还是传感器故障造成质量波动,并提出改进措施。

5 讨论

从生产过程角度考虑,造成中药制剂产品质量波动原因可分为偶然因素和系统因素 2 类。偶然因素造成的质量波动是随机的,称为正常波动。当仅有偶然因素存在时,认为生产过程处于受控状态;系统因素造成的质量波动称为异常波动,当有系统因素存在时,认为生产过程处于失控状态。统计控制图能够区分这 2 种因素,通过检测系统因素造成的质量波动,提示质量管理人员采取措施消除这些异常因素对过程的影响,从而达到提高产品质量的目的。

SPC 可用来促进和保持中药制药企业的健康生产,帮助质量管理人员认识和了解生产程中存在的问题、内在特性和变化规律,并寻找生产过程发生异常的原因。但 SPC 并不能解决中药生产中的所有问题。数据的采集、统计分析计算、图表只是统计过程控制的一部分,统计分析的结果需要质量管理和过程操作人员用经验、知识进行解释,需要根据过程机理去理解、运用。

本文按照建立统计模型的不同阶段,针对清开灵生产过程栀子提取液中栀子苷提取率这一关键质量属性,基于 $\mu \pm 3\sigma$ 原则建立了 Shewhart 控制图,并发展了栀子苷提取率的 EWMA 控制方法。为了对异常过程进行异常原因的追溯,研究了栀子前处理生产工艺参数的 MSPC 控制策略,通过贡献图分

析确定异常来源。USPC 与 MSPC 结合可以更好地理解并监控中药生产过程。

随着过程控制硬件设备的发展,国内越来越多的中药生产企业已建立生产过程集中控制系统,可采集并存储大量的过程测量数据。应用统计过程控制的方法,将隐藏在过程数据中的大量有用信息提炼出来并将其合理应用,对于提高生产的安全性和中药制剂产品质量的均一性和稳定性具有重大意义。然而中药生产过程本身工艺繁杂、设备及控制对象粗糙,随机影响因素众多,变量之间的耦合作用较强,检测数据具有较大的噪声,如何发展并建立起适宜于中药在线生产的精确的统计过程控制模型尚需进一步深入的研究。

[参考文献]

[1] Arawati Agus, Za'faran Hassan. Enhancing production performance and customer performance through total quality management (TQM): Strategies for competitive advantage [J]. Procedia-Social and Behavioral Sci, 2011,24:1650.

[2] 于涛,王高山. 先进制造环境中工序质量统计过程控制的地位与作用[J]. 中国机械工程, 2004, 15 (23):2107.

[3] 王智民,钱忠直,王永炎. 中药产品质量过程控制[J]. 中国中药杂志,2011,36(6):656.

[4] 史新元,张燕玲,王耘,等. 中药生产过程中质量控制的思考[J]. 世界科学技术——中医药现代化, 2008,11(5):121.

[5] M Thomson, P M Twigg, B A Majeed, et al. Statistical process control based fault detection of CHP units[J]. Control Eng Practice, 2000,8 (1):13.

[6] L E Mujicaa, J Vehí, M Ruiza, et al. Multivariate statistics process control for dimensionality reduction in structural assessment[J]. Mech Syst and Signal Proc, 2008,22(1):155.

[7] Hongxia Huang, Haibin Qu. In-line monitoring of alcohol precipitation by near-infrared spectroscopy in conjunction with multivariate batch modeling[J]. Anal Chim Acta, 2011,707(1/2):47.

[8] Hamza Albazzaz, Xue Z Wang, Fatma Marhoon. Multidimensional visualisation for process historical data analysis: a comparative study with multivariate statistical process control[J]. J Process Control,2005,15(3):285.

[9] 唐凯. 基于多元统计过程控制的故障诊断技术[D]. 杭州:浙江大学,2004.

[责任编辑 全燕]