

偏最小二乘法建立三磷酸腺苷二钠注射液近红外快速定量模型

韩莹^{1*}, 曾文珊¹, 周远华¹, 王文佳¹, 赖衍清²[1.广州市药品检验所, 广州 510160; 2.布鲁克(北京)科技有限公司, 北京 100081]

中图分类号 R927 文献标志码 A 文章编号 1001-0408(2015)07-0991-04

DOI 10.6039/j.issn.1001-0408.2015.07.39

摘要 目的:建立可用于现场快速筛查、监督药品质量的三磷酸腺苷二钠(ATP-2Na)注射液的快速定量分析模型。方法:收集全国16家生产企业共73批ATP-2Na注射液样品,采集近红外光谱图,利用化学计量学偏最小二乘法(PLS)将光谱信息与含量数据关联,建立通用型定量分析模型,并进行方法学验证。结果:全部样品经聚类分析分为I类、II类,分开建立两组均质样本定量模型。建立的模型(I)交叉验证均方根误差(RMSECV)为1.73%,相关系数为0.951 2;模型(II)RMSECV为1.55%,相关系数为0.943 1。模型验证试验中,经配对 t 检验,两种模型分别对18批和14批样品的预测值与真实值比较,差异无统计学意义。结论:建立的模型能够快速、准确地预测不同生产企业ATP-2Na注射液中ATP-2Na的含量,并能对不合格样品进行预测,达到现场监督快速测定样品含量的目的。

关键词 近红外光谱;偏最小二乘法;三磷酸腺苷二钠注射液;定量模型

Establishment of Near-infrared Rapid Quantitative Model for Adenosine Triphosphate Disodium Injection by Partial Least Squares

HAN Ying¹, ZENG Wen-shan¹, ZHOU Yuan-hua¹, WANG Wen-jia¹, LAI Yan-qing²[1.Guangzhou Institute for Drug Control, Guangzhou 510160, China; 2.Brook(Beijing) Science and Technology Co., Ltd., Beijing 100081, China]

ABSTRACT OBJECTIVE: To establish the rapid quantitative model for Adenosine triphosphate disodium (ATP-2Na) injection that can be used for on-site rapid screening and monitoring of drug quality. METHODS: 73 batches of ATP-2Na injection were collected from 16 manufacturers to collect their near-infrared (NIR) spectroscopy; the universal quantitative model was established through combing spectral information with content data with chemometrics partial least square (PLS), and then the methodology validation was conducted. RESULTS: According to cluster analysis, all samples were divided into type I and type II, and homogeneous sample quantitative model of 2 groups were established separately. The root mean square error of cross validation (RMSECV) of model (I) and model (II) were 1.73% and 1.55%, respectively; the correlation coefficient of model (I) and model (II) were 0.951 2 and 0.943 1, respectively. According to the paired t -test, there were no statistically significant differences between predicted and measured values in 18 batches and 14 batches of samples in the model validation test. CONCLUSIONS: The model can predict the content of ATP-2Na in ATP-2Na injection from different manufacturers fast and accurately, predict those unqualified samples and achieve the purpose of on-site monitoring and rapid determination of sample content.

KEYWORDS Near-infrared spectrascopy; Partial least square; Adenosine triphosphate disodium injection; Quantitative model

tion parameters on characteristics and internal morphology of poly(D, L-lactide-co-glycolide) microspheres formed by the solvent evaporation method[J]. *Eur J Pharm Biopharm*, 2008, 68(2): 214.

[8] 张海龙.丹皮酚聚乳酸羟基乙酸微球的制备及体外释药考察[J]. *中国药房*, 2013, 24(19): 1765.

[9] Klose D, Siepmann F, Elkharraz K, et al. How porosity and size affect the drug release mechanisms from PLGA-based microparticles[J]. *Int J Pharm*, 2006, 314(2): 198.

[10] Siepmann J, Faisant N, Akiki J, et al. Effect of the size of

biodegradable microparticles on drug release: experiment and theory[J]. *J Control Release*, 2004, 96(1): 123.

[11] Siepmann J, Elkharraz K, Siepmann F, et al. How autocatalysis accelerates drug release from PLGA-based microparticles: a quantitative treatment[J]. *Biomacromolecules*, 2005, 6(4): 2312.

[12] Zhao A, Hunter SK, Rodgers V. Theoretical prediction of induction period from transient pore evolution in polyester-based microparticles[J]. *J Pharm Sci*, 2010, 99(11): 4477.

*主管药师。研究方向:药品快速检验、药物分析。E-mail: han-ningying@126.com

(收稿日期:2014-07-18 修回日期:2014-11-27)

(编辑:邹丽娟)

近红外光谱(Near-infrared spectroscopy)分析技术的发展离不开化学计量学方法的发展,而定量校正(多元校正)作为化学计量学中的一个主要分支,其主要原理就是将样品浓度或其他物理化学性质与分析仪器吸收值之间建立定量关联。偏最小二乘法(Partial least square, PLS)是一种新型的多元统计数据分析方法,在一种算法下,可以同时实现回归建模(多元线性回归)、数据结构简化(主成分分析)以及两组变量之间的相关性分析(典型相关分析)。PLS在近红外光谱定量分析中运用广泛,已逐渐成为一种标准的常用方法^[1-3]。

三磷酸腺苷二钠注射液为辅酶类药,用于进行性肌萎缩、脑出血后遗症、心肌疾患及肝炎等的辅助治疗,市场需求量大。其有效成分三磷酸腺苷本身不稳定,极易受温度、水分等影响而发生水解,从而使其含量下降,影响质量^[4-5]。该产品在药检中主要不合格项目集中在含量测定方面,这可能与该产品在流通过程中贮存温度太高及生产工艺的不同有关。由于法定标准采用紫外法测定总核苷酸含量、高效液相色谱法测定三磷酸腺苷二钠质量比,方法所需检测时间长、操作烦琐,对技术人员要求较高。因此,建立一种快速、有效的含量分析方法在监督药品质量、保证药品疗效方面具有重要意义。笔者利用近红外光谱技术结合PLS算法,以三磷酸腺苷二钠注射液为研究对象,建立均质样本定量分析模型,使其达到可快速测定样品含量从而实现现场监督产品质量的目的。

1 材料

1.1 仪器

Matrix-F型傅里叶变换近红外光谱仪,配有1.5 m长的固体光纤探头及液体测试配件、InGaAs型检测器、OPUS 5.0及7.0光谱分析软件(德国Bruker公司);CARY100型紫外分光光度计(美国Varian公司);1200型高效液相色谱仪,配Kromasil C₁₈分析色谱柱(美国Agilent公司)。

1.2 药品与试剂

国内16个生产企业的73批三磷酸腺苷二钠注射液(规格:2 ml:20 mg)来自2013年国家药品评价抽验品种;其余试剂均为分析纯。

2 方法

2.1 光谱采集条件的确立

2.1.1 温度 温度对液体样品的近红外吸收有较大影响^[6-7],需要在特定温度下建立定量分析模型。根据三磷酸腺苷二钠注射液的贮藏条件(避光,并不超过20℃),因此近红外光谱测定温度控制在(20±3)℃。

2.1.2 光程 选用车载近红外光谱仪液体附件的1号配件(光程:1 mm)、2号配件(光程:2 mm)和5号配件(光程:5 mm),由透反射技术采集样品的近红外光谱。在4 400 cm⁻¹波数处得到吸光度,1号配件测得值为0.434,2号配件测得值为1.5;但5号配件测定时吸光度在4 000~5 000 cm⁻¹处谱段全部溢出,超过仪器测量吸光度的响应范围,这是由于随着光程的增加,水的近红外吸收信号增大,由此将主成分的近红外吸收信息掩盖所致^[8]。

按照分光光度计误差理论,在单波长下,当吸光度为0.434时,对浓度的测定误差最小。因此分析液体样品时,光谱选择

最优的光程可以尽量减少测定误差。因此选用1号配件分析样品的近红外图谱。

2.2 近红外光谱数据采集

溶液温度控制在(20±3)℃,以仪器内置背景为参比;分辨率设定为8 cm⁻¹;背景扫描次数设定为32次;样品扫描次数设定为32次;扫描范围为12 000~4 000 cm⁻¹。每批次样品取6支进行样品测试,得到6张光谱。建立定量分析模型时,取6支样品光谱的平均光谱用于建模。

2.3 高效液相色谱分析

所有样品均按三磷酸腺苷二钠注射液国家药典标准方法^[9]测定,得到的样品含量范围为8.49~10.95 mg/ml,按C₁₀H₁₄O₁₃N₅P₃Na₂计算。

3 结果与分析

3.1 建立均质定量分析模型

3.1.1 训练集、验证集样本的划分 对73批样品进行聚类分析,选取全谱段与矢量归一预处理,结果见图1。

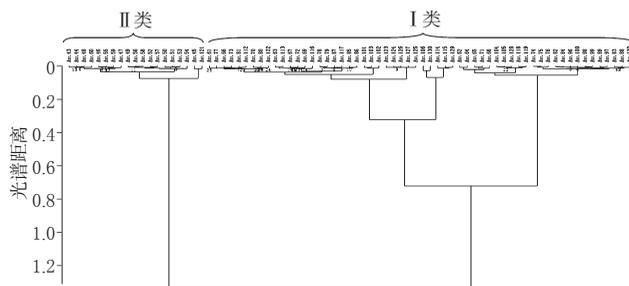


图1 总样本聚类分析树形图

Fig 1 Hierarchical cluster analysis for samples

由图1可以看出,全部样品聚类分析分为I类、II类。建立模型时,先以I类均质样本中的59批样品,建立均质模型(I);再将第II类(包含14批样品)均质样本加入模型,建立通用型模型。若建立的通用型模型预测误差较大,则需要分别建立I类样本和II类样本的均质定量模型,以保证模型能够准确预测结果^[10]。

建立均质定量模型(I):将59批图谱相近的样品聚类分析,结合质量浓度梯度分布,选出41批样品(质量浓度范围为8.49~10.95 mg/ml)作为训练集样本,余下18个(质量浓度范围为8.78~10.81 mg/ml)样品作为验证集样本,得到训练集样品和验证集样品的第一主成分对第二主成分得分散点分布图(见图2)。结果训练集样品或验证集样品分布都较为均匀,表明训练集样本具有足够的代表性。

3.1.2 特征谱段的选择 比较三磷酸腺苷二钠注射液(样品质量浓度:10.27 mg/ml)和实验室用纯水的近红外光谱,见图3。

由图3可见,5 500 cm⁻¹和6 800 cm⁻¹波数附近的强吸收谱带为水的吸收峰。将三磷酸腺苷二钠注射液分别稀释成3.4、5.1、6.8 mg/ml的溶液,扫描图谱后,以矢量归一化进行数据预处理,发现在波数6 400~5 400 cm⁻¹和4 700~4 200 cm⁻¹范围,样品中组分质量浓度与吸收度关系具有相关性,见图4。表明这两个谱段或许反映了样品中组分的近红外吸收信息。因此,以6 100~5 400 cm⁻¹和4 700~4 200 cm⁻¹两个谱段为建模的基本谱段。

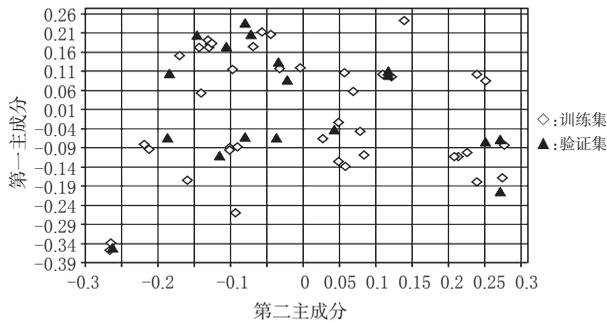


图2 第一主成分对第二主成分得分散点分布图

Fig 2 Scatter distribution diagrams of the first principal component vs. the second principal component

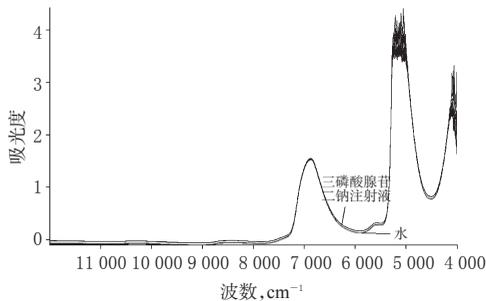


图3 三磷酸腺苷二钠注射液和纯水的近红外光谱图比较

Fig 3 Comparison of NIR spectra diagrams between ATP-2Na injection and pure water

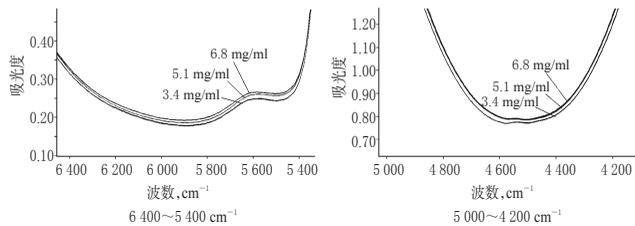


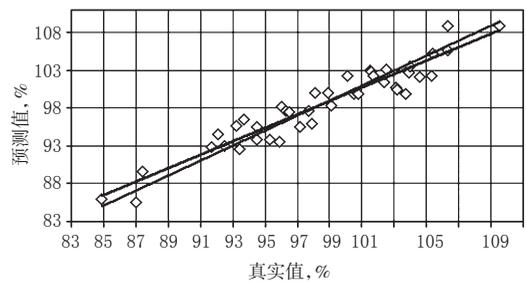
图4 不同质量浓度三磷酸腺苷二钠注射液局部近红外光谱图

Fig 4 Local NIR spectra diagrams of ATP-2Na injection with different concentrations

3.1.3 图谱预处理方法的选择 在OPUS软件“建立定量2方法”程序中,提供了12种预处理方法。取6 100~5 400 cm^{-1} 和4 700~4 200 cm^{-1} 两个谱段范围,组合12种数据预处理方法,采用软件自动优化建模参数,建立均质样本(I)的定量分析模型。优化结果显示,选择矢量归一化时,所建模型的内部验证和外部验证的均方差均较小,因此以矢量归一化法进行数据预处理。

3.1.4 建立均质样本(I)定量模型 对训练集样本进行内部交叉验证,以内部验证均方差(RMSECV)为优化参数的指标,得到均质样本(I)模型参数,谱段范围为:6 395.3~5 361.6 cm^{-1} 和4 852.4~4 243 cm^{-1} ,预处理:矢量归一化。结果,模型的RMSECV为1.73%,检验集验证均方根误差(RMSEP)为1.79%,相关系数(r)为0.951 2,主成分数即维数(Rank)为8。预测值与真实值的线性关系图见图5,表明样品预测值与真实值有很好的相关性。

3.1.5 对均质样本(II)的预测 利用已建立的定量模型(I)



偏移量:10.660 斜率:0.891 相关系数:0.951 2
维数:8 $R^2=90.46$ RMSECV=1.73% 偏移:0.037 5

图5 预测值与真实值的线性关系示意图

Fig 5 Linear correlation diagram of predicted value and measured value

对II类均质样本进行含量预测。结果预测值与真实值的绝对误差在25~61 mg/ml内,相对误差在23%~47%内,二者有一定差距,表明均质样本(I)的定量模型并不适用于均质样本(II)的含量预测。故考虑将部分均质样本(II)的光谱加至定量模型(I)的训练集中,重新优化参数,建立通用型定量分析模型。

3.2 尝试建立通用定量模型

在14批均质样本(II)中选取10批样品,加入到均质定量模型(I)中的训练集;并将余下的4批加入到验证集中,重新优化建模条件,得到三磷酸腺苷二钠注射液通用定量模型。参数设置:谱段范围为5 878.2~5 361.4 cm^{-1} ,光谱预处理方法为二阶导数(17点平滑),维数为3, RMSECV为4.12%。以验证集的22批样品进行外部验证,得到RMSEP为4.47%,真实值与预测值相关系数为0.600,表明两者相关性较差,建议分开建立模型。

3.3 建立均质样本(II)定量模型

由于样本量小,以均质样本(II)14批样品(质量浓度范围为9.09~10.79 mg/ml)全部作为训练集,内部交叉验证的方法建立模型。优化后的均质样本(II)模型参数:谱段范围为4 825.4~4 243 cm^{-1} ,预处理为一阶导数(17点平滑)。得到模型的RMSECV为1.55%,相关系数为0.943 1,维数为5。

3.4 模型验证

3.4.1 专属性 利用马氏距离与限定值的比较来评价模型的专属性^[11],模型的限定值=Factor \times Rank/M。式中,Factor为因子,一般因子范围为2~10,本试验选择为2;Rank为主成分维数,本试验选择为5;M为建模样品数,本试验选择为44。故均质模型(I)、(II)的限定值分别为0.39、0.71。利用所建模型对与三磷酸腺苷二钠具有同样母核的阿糖腺苷制备成的注射液进行识别,结果样品的马氏距离明显大于模型限定值。另取辅料按企业的处方制备成不含成分的阴性样品,利用模型识别,结果阴性样品的马氏距离均大于限定值。由此可见,所建模型对三磷酸腺苷二钠注射液具有良好的专属性。

3.4.2 线性关系 由待分析物的预测结果与参比值之间的关系来评价。均质模型(I):在8.49~10.95 mg/ml范围内,以近红外预测值(y)和参比值(x)求得线性方程为 $y=8.91 \times 10^{-1}x+0.375 \times 10^{-1}$ ($r=0.951 2$);均质模型(II):在9.09~10.79 mg/ml

范围内得线性方程为 $y = 8.34 \times 10^{-1}x - 0.0191 \times 10^{-1}$ ($r = 0.9431$), 两模型线性关系均良好。

3.4.3 精密度 取样品在同一天内、由同一操作者在同样的操作环境中利用均质模型(I)、(II)预测其主成分的含量,共预测10次,其RSD分别为1.22%和1.34%,可见两模型精密度均良好。

3.4.4 准确性 定量模型(I)对验证集18批样品的含量预测结果见表1。

表1 定量模型(I)对验证集的预测结果

Tab 1 Predicted results of validation set for model(I)

序号	文件名	真实值, %	预测值, %	偏差, %
1	Av.100	87.8	88.5	-0.68
2	Av.101	94.7	94.2	0.49
3	Av.115	98.6	100.1	-1.48
4	Av.117	99.0	100.2	-1.15
5	Av.126	100.6	98.3	2.29
6	Av.457	108.1	107.1	1.01
7	Av.460	102.9	103.5	-0.59
8	Av.462	102.4	105.4	-2.99
9	Av.465	104.0	103.9	0.07
10	Av.468	104.6	103.7	0.88
11	Av.62	101.0	98.1	2.86
12	Av.63	94.8	96.7	-1.86
13	Av.64	96.3	95.8	0.46
14	Av.67	95.6	95.5	0.08
15	Av.74	99.7	98.0	1.74
16	Av.82	103.4	99.1	4.28
17	Av.84	102.0	99.5	2.53
18	Av.86	97.2	98.4	-1.17

由表1可见,预测结果的平均偏差为1.48%,最大为4.28%;另均质模型(II)定量模型对训练集14批样品的含量预测结果的平均偏差为1.25%,最大为3.03%(具体数据略)。利用配对t检验,在95%置信区间内,两模型预测值与参比值比较,差异无统计学意义。

4 讨论

4.1 温度对液体样品近红外光谱的影响

温度对液体样品近红外光谱的影响较大,在建立液体定性定量模型时,为了保证图谱的稳定性和模型的准确性,必须选择合适的方法和途径消除温度的影响。另外,研究表明,温度变化对液体样品的影响是有规律可循的,通过建立校正曲线可以校正结果,实现对样品的准确定量分析^[11]。

4.2 活性成分与水分子的相互作用

水溶液样品中的活性成分在近红外的吸收相对于水的吸收较弱。求取样品近红外光谱对水近红外光谱的第二因子谱,可以反映样品的吸收情况。选择全谱段(12 000~4 000 cm^{-1}),预处理方法为矢量归一化,求取均质样本(I)和均质样本(II)样品的平均光谱对水的第二因子谱,样品光谱在5 000~4 200 cm^{-1} 谱段存在差异,而且在水的吸收谱段,即8 000~5 400 cm^{-1} 谱段也体现出一定差异,这就提示主成分水

溶液的近红外吸收情况不仅与主成分的吸收特征有关,而且与水分子的相互作用也有关系。

4.3 对不合格样品的预测

三磷酸腺苷二钠注射液均质模型(I)质量浓度范围为8.49~10.95 mg/ml,能够对不合格样品(质量浓度低于标示值的90%)进行准确预测(见表1)。由于模型覆盖的质量浓度范围应该宽于预测样品的范围,因此需要收集更多的样品及制备更多质量浓度样品来拓展模型含量预测范围,进一步提升模型的应用性。

4.4 适用范围及注意事项

近红外光谱法定量分析依赖于校正模型,而校正模型在实际应用中需要技术人员不断维护,因此该方法不适合于零散样本的分析。同时要求光谱仪应具有长期的稳定性,不仅仪器的各项性能指标不能发生显著改变,而且光谱仪光路中任何一个光学部件的更换,都可能会使模型失效。

参考文献

- [1] 孙美玲,相秉仁.近红外透射光谱法快速无损定量分析甲磺酸加替沙星注射液中的主成分[J].中国药科大学学报,2003,34(3):46.
- [2] 李伟,孙素琴,覃洁萍,等.近红外漫反射法测定杜仲中松脂醇二葡萄糖苷的含量[J].中国中药杂志,2010,35(24):3 318.
- [3] 陈久艳,方敏,吴再珍,等.基于近红外光谱测定复方盐酸阿米洛利片有效成分的含量[J].中国药房,2013,24(9):828.
- [4] 金睿洁.温度对三磷酸腺苷二钠注射液含量的影响[J].中国当代医药,2012,5(12):139.
- [5] 郭丙炎,吴胜丹,王岚,等.三磷酸腺苷二钠及其制剂含量测定方法的改进[J].中国药品标准,2002,3(5):28.
- [6] 陈韵,史振志,徐可欣,等.溶液近红外光谱分析中温度影响的修正[J].光谱学与光谱分析,2009,29(11):2 966.
- [7] 褚小立,王艳斌,陆婉珍.近红外光谱定量校正模型的建立及应用[J].理化检验:化学分册,2008,46(8):796.
- [8] 杜敏,吴志生,林兆洲,等.光程对清开灵注射液中黄芩苷近红外定量模型的影响[J].药物分析杂志,2012,32(10):1 796.
- [9] 国家药典委员会.中华人民共和国药典:二部[S].2010年版.北京:中国医药科技出版社,2010:29.
- [10] 卢京光,王立云,李新荣,等.近红外光谱法快速分析硫普罗宁注射液的含量[J].中国药学杂志,2012,47(16):1 328.
- [11] 胡昌勤,冯艳春.近红外光谱法快速分析药品[M].北京:化学工业出版社,2010:155-158.

(收稿日期:2014-07-09 修回日期:2014-08-15)

(编辑:刘 萍)