

近红外漫反射光谱法对阿莫西林胶囊的定量及定性分析

沈漪¹ 潘颖¹ 刘全² 项竞佐¹

(1.上海市药品检验所 上海 200233; 2.美国 Thermo Nicolet 公司)

摘要 目的: 采用近红外漫反射光谱分析技术对阿莫西林胶囊进行定量和定性分析。**方法:** 按阿莫西林胶囊配方组成配制含主药阿莫西林浓度范围从 5.91%~84.13%的 32 个实验室样品,并收集来源于九个厂家的 41 批工业样品,采集其近红外光谱。分别采用偏最小二乘回归和判别分析法建立定量和定性分析模型,将其用于对未知样品含量进行预测分析,并对定量分析方法的重现性和加样回收率进行考察。**结果:** 定量分析模型对 24 个未知样品的预测均方差 RMSEP 为 1.22%,预测值与真值的相关系数 R 为 0.9983,加样平均回收率为 99.75%,系统精密度 RSD 为 0.3%,方法精密度 RSD 为 0.6%;定性分析模型对 18 个预测样品的判错率为 0%。**结论:** 用近红外光谱分析技术对阿莫西林胶囊进行定量和定性分析结果准确可靠,方法简便快速,不需预处理,可推广用于此类样品工业现场的原位和在线检测。

关键词 近红外 漫反射光谱 阿莫西林 定量分析

Determination and qualification of the Amoxicillin Capsules by

NIR Diffuse Reflectance Spectroscopy

Shen Yi¹, Pan ying¹, Liu Quan², Xiang Jingzuo¹

(1. Shanghai Institute for Drug Control, Shanghai 200233)

(2. Thermo Nicolet corporation, U.S.A)

Abstract Objective To determinate and qualification of the amoxicillin capsule by near-infrared (NIR) Diffuse Reflectance Spectroscopy. **Method** We prepared 32 samples which concentration range of amoxicillin is from 5.91% to 84.13%, at the same time collect 41 batches samples which came from 9 factories, then use Partial least square(PLS) to build the calibration model to predict the concentration of unknown sample, and examine the method reproducibility and recovery. We use discriminant analysis to build classification model.**Result** A model was set up using a calibration set (73 samples). The determination coefficient R is 0.9994, which came from cross-validation, and root mean square error of cross-validation (RMSECV) is 0.96. Root mean square error of prediction (RMSEP) obtained from the validation including 24 unknown samples is 1.22%, the correlation coefficient between the true value and prediction value is 0.9983. The average recovery is 100.25%, RSD of the system precision is 0.3%, RSD of the method precision is 0.6%. The misclassified rate is 0% on the qualification model using discriminant analysis. **Conclusion** The results show that the predict model using NIR is reliable, the method is fast and exact, and can be applied to on-line detection.

Key words near infrared, diffuse reflectance spectrophotometry, amoxicillin, quantitative analysis

近红外(Near Infrared, NIR)光谱是指处于可见光和中红外光之间,波长范围为 780 - 2500 nm (12000 - 4000 cm^{-1})的一段谱区,其光谱信息主要反映的是有机物分子中C-H、N-H、O-H等含氢基团的倍频与合频吸收^[1]。由于谱带严重重叠,谱峰难以分辨,该谱区曾一度被称为“光谱技术中的沉睡者”^[2]。近几年来,随着化学计量学技术的发展及先进光谱仪器的出现,人们重新认识了它的价值。目前,NIR光谱分析技术因其快速、无损、多成分同时测定等特性,已被广泛应用于农业^[3]、石油^[4]、化工^[5]等行业,而在医药领域,它已被成功用于包括原辅料^[6]、制剂^[7,8]、生产过程^[9]等几乎所有环节的定性和定量分析。

阿莫西林（Amoxicillin）胶囊是一种口服半合成青霉素，具有抗菌谱广，杀菌力强，毒性低等特点，适用于轻中度感染，已广泛应用于临床。目前其含量测定方法为高效液相色谱法，该方法需进行复杂的样品制备和预处理工作，不能用于工业现场的原位快速分析。

本文以阿莫西林胶囊为例，按一个厂家的制剂配方配制了 32 个样品，同时收集了九个厂家的阿莫西林胶囊制剂共 41 批，采集其近红外光谱，分别应用偏最小二乘回归算法(PLS)和判别分析技术建立了定量和定性分析模型，并对未知样品进行预测分析。结果证明，该方法准确可靠，可推广用于此类样品工业现场的定性和定量快速检测。

1 仪器与样品

Antaris 傅立叶变换 NIR 光谱仪（美国 Thermo Nicolet 公司），配有积分球漫反射采样系统，Result 操作软件，TQ Analyst 6.2 分析软件；

阿莫西林原料（纯度 86.74%），辅料硬脂酸镁、淀粉为药用规格。按阿莫西林胶囊配方组成配制含主药阿莫西林浓度范围从 5.91%~84.13% 的 32 个实验室样品。

九个厂家的阿莫西林胶囊制剂共 41 批，分别是：昆明贝克诺顿制药有限公司 4 批，上海衡山药业有限公司 8 批，香港澳美制药厂 5 批，香港联邦制药厂有限公司 4 批，丽珠集团丽珠制药厂 4 批，浙江金华康恩贝生物制药有限公司 4 批，甘肃兰药药业集团有限责任公司 4 批，海南三叶制药厂有限公司 4 批，山东鲁抗医药股份有限公司 4 批。

2 NIR 光谱采集

将样品分别装于样品瓶中，在 $10000 \sim 4000 \text{ cm}^{-1}$ 波数范围内采集其 NIR 光谱，扫描次数 40 次，分辨率 8 cm^{-1} ，每个样品重复 3 次，取平均值。图 1 为阿莫西林和辅料的 NIR 光谱图。图 2 和图 3 分别为 32 个实验室样品及 41 个工业样品的原始光谱及一阶导数光谱。

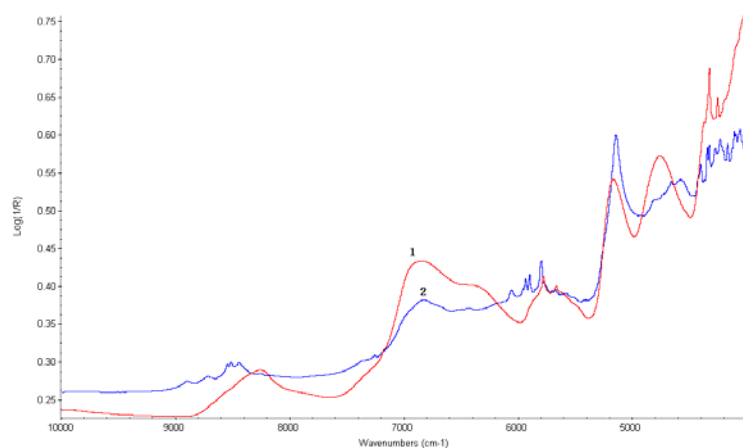


图 1 阿莫西林和辅料的 NIR 光谱图

Fig.1 NIR spectra of excipient and amoxicillin

1.辅料（excipient） 2.阿莫西林（amoxicillin）

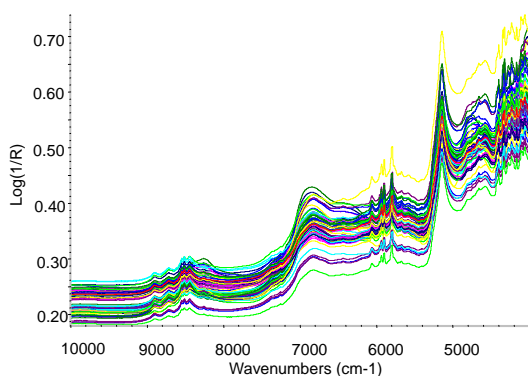


图2 样品的原始 NIR 漫反射光谱图

Fig.2 Original NIR spectra

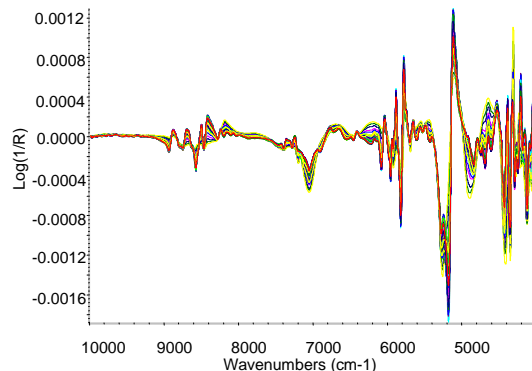


图3 一阶导数光谱图

Fig.3 First derivative spectra

3 定性分析

将九个厂家 41 批样品的 NIR 光谱图用判别分析法建立定性分析模型。采用一阶导数，并选取光谱波数范围为 $9881.5-4119.2\text{cm}^{-1}$ ，作为建模参数，并用建立的模型对其中 18 个未知样品进行预测，以验证模型的预测效果，结果判错数为 0，证明该模型能很好的对建模的九个厂方的样品进行识别。其优化过程如下：

3.1 光谱范围及数据处理方法的选择

尝试用三个一阶导数光谱区别较明显的三个谱段： $9221.6-7965.9\text{cm}^{-1}$ 、 $7965.0-6173.3\text{cm}^{-1}$ 和 $6173.0-4000.0\text{cm}^{-1}$ ，以及仪器推荐的光谱范围 $9881.5-4119.2\text{cm}^{-1}$ （选取导数光谱时均采用 Norris 平滑滤波进行校正）。表 1 为采用不同光谱范围和光谱预处理方法所建立的判别分析模型的误判样品数比较。

表 1 采用不同光谱范围和光谱预处理方法所建立的判别分析模型的误判样品数比较

Table1 Influence of NIR spectra range and pretreatment method on the number of misclassified samples

序号 No.	波数范围 (cm^{-1}) Wavenumber range (cm^{-1})	原始光谱 Original NIR spectra		一阶导数光谱 First derivative spectra		二阶导数光谱 Second derivative spectra	
		校正集	验证集	校正集	验证集	校正集	验证集
1	9881.5 - 4119.2	2	0	0	0	0	0
2	9221.6 - 7965.9	18	3	0	0	1	1
3	7965.0 - 6173.3	7	1	0	0	0	2
4	6173.0 - 4000.0	4	0	0	0	0	0

从表 1 可见用一阶导数光谱在任何光谱范围的判错数均为 0，用原始光谱的判错数最高，二阶导数光谱其次，说明采用一阶导数光谱可以非常好的消除基线的偏移，扣除本底，使各个不同厂家的特征区别更明显，从而达到良好的区分；另外，尽管在四个光谱范围一阶导数的判错率均为 0%，但考虑到由于厂家较多，选取光谱范围越宽，利用的信息越充分，模型更稳健，最终选取波数范围为 $9881.5-4119.2\text{cm}^{-1}$ 。

3.2 定性分析模型及样品判别结果

主成分分析（Principle Component Analysis, PCA）是最常用的判别分析方法⁽⁷⁾，它可以将多波长下的光谱数据压缩到有限的几个因子空间内，再通过样品在各因子空间的得分确定所属的类别。光谱数据经前述一节导数处理并选择光谱范围后，对其作 PCA 分析，以第一主成分对第二主成分作图，如图 4。

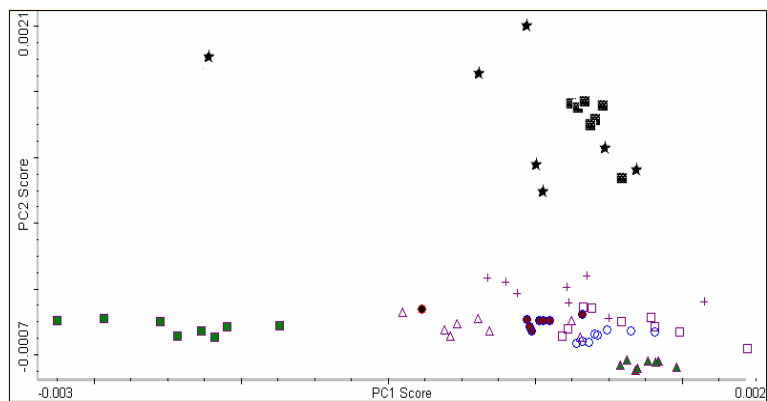


图 4 样品在前两个主成分中的得分分布

Fig. 4 Scatter plot of scores for the first and second PC of the 41 samples from 9 different factories

其中★香港澳美制药厂 ▲山东鲁抗医药股份有限公司 ●浙江金华康恩贝生物制药有限公司 ○海南三叶制药厂有限公司 ■甘肃兰药药业集团有限责任公司 △昆明贝克诺顿制药有限公司 □上海衡山药业有限公司 ▨丽珠集团丽珠制药厂 + 香港联邦制药厂有限公司

主因子数的选择是影响 PCA 法效果的一个关键因素。主因子数太少，会导致信息利用率过低，主因子数太多，则会引入许多无用或噪音信息，两者均会降低模型预测精度和稳健性。在本文的实验中，主成分数的选择对模型判错率的影响情况如表 2。从中可见，当主因子数为 15 时，即可以对 9 个厂家的样品进行准确地判别分析，此时，光谱信息的利用率为 99.93%。

表 2 主成分数对模型判错率及对累计贡献的影响

Table 2 The infection of principle components to misclassified and cumulative contribution

编号	主成分数	累计贡献%	判错数
1	10	99.74	3
2	14	99.91	1
3	15	99.93	0
4	16	99.94	0
5	20	99.96	0

4 定量分析

4.1 数据处理方法

分别从 32 个实验室样品和 41 个工业样品中随机选取 49 个样品组成校正集，用于建立定量分析模型，其余 24 个样品当作未知样品组成预测集，用于验证模型的预测效果。光谱数据经多元散射校正（Multiplicative Scatter Correction, MSC）预处理和 Norris 导数平滑滤波后，采用偏最小二乘（Partial Least Square, PLS）回归法建立定量校正模型，以内部交叉验证均方差 (RMSECV) 为指标优化建模参数，以未知样品的预测均方差 (RMSEP) 及其相对偏差 (RSEP) 考察模型的预测准确性，RMSECV、RSECV、RMSEP 和 RSEP 的计算公式见文献^[10]。

4.2 定量分析模型的建立

4.2.1 光谱预处理方法的选择

表 3 为分别采用原始光谱、一阶导数光谱和二阶导数光谱进行建模时的 RMSECV 值。从表 3 可见采用原始光谱时的 RMSECV 最小，这同建立定性分析模型时的光谱预处理方法

不同，因为样品的定性分析除了涉及样品组成信息外，还涉及到制剂工艺的不同带来的样品物理性质的差异，用一阶导数光谱消除各方面的影响选取特征信息来进行区分类别，而定量往往只需考虑组成浓度的不同，在本实验中则只需原始光谱即能达到最优化。

表 3 导数处理对校正模型 RMSECV 的影响*

Table 3 RMSECV result of model using derivative spectra and original spectra

No.	光谱预处理方法	主因子数	RMSECV
1	原始光谱	8	0.96
2	一阶导数	10	1.05
3	二阶导数	6	1.48

*光谱范围：10000-4000 cm^{-1}

4.2.2 光谱范围的选择

表 4 为使用不同谱区的光谱：全谱带 10000 - 4000 cm^{-1} 、软件推荐光谱带 8458.2 - 8331.0、8523.8 - 8489.1、8743.7 - 8651.1 cm^{-1} 与自选光谱带 9221.6 - 7965.9 cm^{-1} 、7965.0 - 6173.3 cm^{-1} 、6173.3 - 4000.0 cm^{-1} 信息建立模型时的内部交叉验证均方差（RMSECV）值，其结果表明在 10000 - 4000 cm^{-1} 范围内内部交叉验证均方差（RMSECV）最小，证明其最佳谱区为 10000 - 4000 cm^{-1} 。

表 4 不同光谱范围的选择对参数的影响*

Table 4 Effect of parameters on various wavenumbers of the spectra

No.	光谱范围 (cm^{-1})	主因子数	RMSECV
1	10000-4000	8	0.96
2	8458.2-8331.0,8523.8-8489.1,8743.7-8651.1	7	1.43
3	9221.6-7965.9	9	1.29
4	7965.0-6173.3	8	1.45
5	6173.3-4000.0	9	2.37

*光谱预处理方法：MSC+平滑+原始光谱

基于前述优化得到的光谱预处理方法、光谱范围及主因子数，即可建立定量分析模型。

4.3 样品分析

用前面所建立的定量校正模型对验证集中 24 个样品浓度进行预测，预测值与真值的相关图见图 4。从图中可见，24 个未知样品的 NIR 预测值能够很好地逼近实际值，相关系数达到了 0.9983，RMSEP、RSEP 分别为 1.22%、0.96%。

表 5 为选取各个厂家的样品用 HPLC 测得的真实值与用模型预测值的比较。

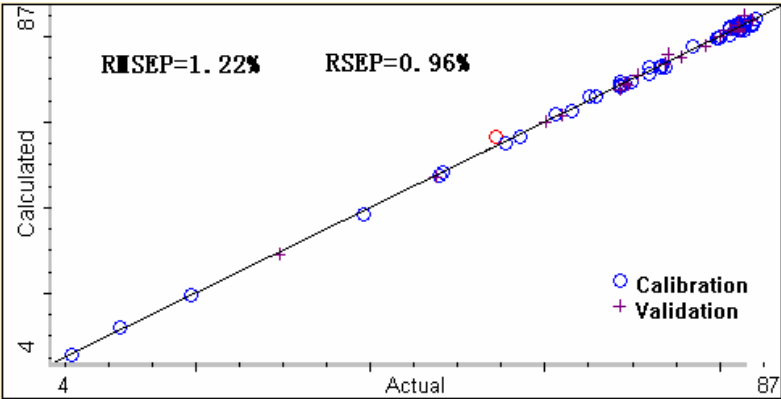


图 5 校正集样品与预测样品真实值与预测值相关图

Fig.5 Scatter plots showing correlation between actual value and predicted value on the calibrated model

表 5 各个厂家的样品真实值与预测值的比较

Table 5 Comparison between actual value and predicted value on the calibrated model

编号	生产厂家	真实值	预测值	相对偏差%
1	香港澳美制药厂	74.19	75.68	2.01
2	昆明贝克诺顿制药有限公司	82.8	84.89	2.52
3	上海衡山药业有限公司	81.50	81.49	0.01
4	浙江金华康恩贝生物制药有限公司	82.05	82.82	0.94
5	香港联邦制药厂有限公司	80.02	79.88	0.17
6	山东鲁抗医药股份有限公司	83.74	83.87	0.16
7	甘肃兰药药业集团有限责任公司	81.22	81.69	0.58
8	丽珠集团丽珠制药厂	68.68	67.71	1.41
9	海南三叶制药厂有限公司	82.18	81.34	1.02

由表 5 可见，九个厂家的样品中相对偏差最大为 2.52%，最小为 0.01%，说明该模型可以对来自这九个厂家的样品的含量进行准确的测定。

4.4 加样回收率

选取上海衡山药业有限公司提供的其中一批样品加入原料适量，按第 2 节方法采集各样品 NIR 光谱，利用前述的定量校正模型测试其浓度，计算加样回收率，结果见表 4。加样回收率范围为 97.66%~101.80%，平均回收率为 99.75%，RSD 为 1.4%。

表 6. 加样回收率结果

Table 6 Result of recovery test

编号	样品浓度%	测得浓度%	回收率%
1	91.58	89.44	97.66
2	92.36	93.07	100.77
3	90.78	90.05	99.20
4	91.24	90.04	98.69
5	92.17	93.83	101.80
6	92.68	93.01	100.36
平均回收率			99.75
RSD(%)			1.4%

4.5 系统精密度与方法精密度

选取配制的 14 号样品，重复测试 6 次，测得系统精密度 RSD 为 0.3%；

选取配制的 14 号样品，装于不同的测试样品瓶测定 6 次，得方法精密度 RSD 为 0.6%。

5. 结论

1. 本文应用 NIR 漫反射光谱分析技术成功地对阿莫西林胶囊制剂进行了同时定性和定量分析，该方法准确可靠、方便快捷，定量分析系统精密度、方法精密度以及加样回收率均良好，显示出良好的实用价值和应用前景。
2. NIR 光谱定量分析模型的建立对校正集样品化学组成的分布范围有一定要求，但实际上，工业上收集到的制剂样品化学组成范围往往都很小，为此，本文采用了工业样品和实验室配置样品相结合的方法，以扩大校正集样品化学组成的分布范围，取得了良好效果。
3. 值得注意的是，近红外光谱信息中不仅包含样品的化学组成信息，同时还与样品的物理性质有关。如果仅采用实验室配置的样品建立定量分析模型，我们发现所建立的模型难以适用于不同厂家的工业样品的预测分析。这是由于各厂家原辅料及制剂的工艺条件的不同带来了样品除了化学组成之外的一些物化性质的差异，这严重限制了实验室配置样品建立的分析模型的适用性。本文采用的实验室配置样品与工业样品相杂交的方法也同

时很好地解决了这一问题。

参考文献:

1. 瞿海斌, 刘全, 程翼宇. 近红外漫反射光谱法测定黄连浸膏粉中生物碱含量, *分析化学*, 已录用。
2. D.L. Wetzel, *Anal. Chem.* 1983, 55: 1165A
3. Maurizio Ventura, Anton de Jager, Herman de Putter, Frans P.M.M. Roelofs. Non-destructive determination of soluble solids in apple fruit by near infrared spectroscopy (NIRS), *Postharvest Biology and Technology*, 1998,14: 21–27
4. Macho S., Larrechi M.S., Near-infrared spectroscopy and multivariate calibration for the quantitative determination of certain properties in the petrochemical industry, *trends in analytical chemistry*, 2002, 21(12):799-806
5. J. Amador-Hernández, J.M. Fernández-Romero, M.D. Luque de Castro. Near infrared thermal lens spectrometry for the real-time monitoring of supercritical fluid extraction, *Talanta*, 1999,49: 813–823
6. A. Candolfi, R. De Maesschalck. Identification of pharmaceutical excipients using NIR spectroscopy and SIMCA, *J. Pharm. Biomed. Anal.* , 1999,19: 923-935
7. M.Blanco, J. Coello, H.Iturriaga. Development and validation of a near infrared method for the analytical control of a pharmaceutical preparation in three steps of the manufacturing process[J], *Fresenius. J. Anal. Chem.*,2000, 368: 534-539
8. Xiangji Zhou, Patricia Hines. Matthew W. Borer, Moisture determination in hygroscopic drug substances by near infrared spectroscopy[J], *J. Pharm. and Biomed. Anal.*, 1998,17: 219-225
9. John D. Kirsch, James K. Drennen. Determination of film-coated tablet parameters by near-infrared spectroscopy[J], *J. Pharm. Biomed. Anal.*, 1995,13: 1273-1281
10. Blanco M., Coello J., Iturriaga H.. *Chemometrics and Intelligent Laboratory Systems*, 2000, 50: 75~82