

化学模式识别方法在中药质量控制研究中的应用进展

孙立丽¹, 王 萌², 任晓亮^{1*}

1. 天津中医药大学中药学院, 天津 300193

2. 天津中医药大学中医药研究院, 天津 300193

摘 要: 化学计量学是以计算机和近代计算技术为基础的一门新兴交叉学科, 在中药鉴别、定性表征、质量控制、组效关系等研究中均具有广泛应用, 尤其在中药的质量控制与评价研究中具有重要意义。综述近年来化学计量学中化学模式识别方法, 包括2种无监督模式识别方法(聚类分析、主成分分析)和4种有监督模式识别方法(簇类独立软模式法、偏最小二乘法判别分析、支持向量机、人工神经网络), 并从产地、基原、炮制、真伪等多个方面总结了化学模式识别方法在中药质量控制研究中的应用。

关键词: 化学模式识别; 化学计量学; 质量控制; 中药; 聚类分析; 主成分分析; 簇类独立软模式法; 偏最小二乘法判别分析; 支持向量机; 人工神经网络

中图分类号: R286.01 文献标志码: A 文章编号: 0253-2670(2017)20-4339-07

DOI: 10.7501/j.issn.0253-2670.2017.20.031

Application progress on chemical pattern recognition in quality control of Chinese materia medica

SUN Li-li¹, WANG Meng², REN Xiao-liang¹

1. School of Chinese Materia Medica, Tianjin University of Traditional Chinese Medicine, Tianjin 300193, China

2. Tianjin State Key Laboratory of Modern Chinese Medicine, Tianjin University of Traditional Chinese Medicine, Tianjin 300193, China

Abstract: Chemometrics is a new cross discipline based on computer and modern technology. It has been widely used in the research of Chinese materia medica (CMM) identification, qualitative characterization, quality control, and group-effect relationship, especially in quality control and evaluation of CMM. In this paper, the application and progress of chemical pattern recognition methods in chemometrics for quality control of CMM in recent years are reviewed. Two unsupervised pattern recognition methods (cluster analysis and principal component analysis) and four supervised pattern recognition methods (soft independent modeling of class analogy, partial least-squares discriminant analysis, support vector machine, and artificial neural network) are described. This paper reviews application of chemical pattern recognition in quality control of CMM from different aspects, including growing areas, herbal origin, processing, identification of the authenticity, etc.

Key words: chemical pattern recognition; chemometrics; quality control; Chinese materia medica; cluster analysis; principal component analysis; soft independent modeling of class analogy; partial least-squares discriminant analysis; support vector machine; artificial neural network

中药具有系统的科学理论、独特的辨证论治思想, 历经了数千年的发展和应用, 蕴含着深厚的人文底蕴和科学内涵, 在中医药理论指导下引入现代科学、实现中药现代化是中药发展的必由之路^[1]。中药质量控制与评价是中药现代化研究的关键问题

之一, 是实现中药标准化的重要基础, 然而中药本身化学成分复杂, 同时基原品种、产地来源、生长环境、采收季节、加工炮制工艺等诸多因素进一步增加了中药质量控制与评价研究的难度。随着现代科学技术特别是仪器分析、计算机科学、化学计量

收稿日期: 2017-06-20

基金项目: 国家自然科学基金资助项目(81473543)

作者简介: 孙立丽, 博士在读, 中药学专业。Tel: 18322050489 E-mail: 18322050489@163.com

*通信作者 任晓亮, 副教授, 主要从事中药分析研究工作。Tel: (022)59596221 E-mail: xiaoliang_ren@sina.com

学等学科的迅速发展, 中药质量控制与评价研究的新思路与新方法不断出现, 有效地促进了中药标准化研究。针对中药本身的复杂性及特有的药性理论基础, 在中医药理论指导下寻找一种适用于中药特点的质量控制与评价方法, 解决中药质量控制的数据化、标准化难题, 并在此基础上用科学的语言阐述中药的基础理论是中药现代化亟待解决的关键问题之一。为完善中药质量标准体系, 促进中医药产业的健康发展, 刘昌孝院士^[2-3]提出中药质量标志物(Q-markers)的这一新概念, 为规范中药质量的研究和标准的建立奠定了基础, 有利于中药全程质量控制和质量溯源体系的建立。

化学计量学是以计算机和近代计算技术为基础的一门新兴交叉学科, 通过统计学或数学方法在化学体系的测量值与体系的状态之间建立联系, 在中药鉴别、定性表征、质量控制、组效关系等研究中均有广泛应用, 尤其在中药的质量控制与评价研究中具有重要意义^[4]。化学模式识别是化学计量学的重要组成部分, 也是筛选中药 Q-markers 的重要数学方法, 按照有无训练可划分为无监督的模式识别和有监督的模式识别方法。前者指在无样品类别信息的情况下, 进行学习或训练, 获取分类信息的方法; 后者则根据样品特征和已知类别的样品(训练集), 用特定的方法或模型进行学习或训练, 从而建立分类模型, 再根据获取的分类模型和未知样品的特征, 对未知样品进行分类^[5]。其中无监督的模式识别方法包括聚类分析(cluster analysis, CA)、主成分分析(principal component analysis, PCA)等。有监督的模式识别方法包括簇类独立软模式法(soft independent modeling of class analogy, SIMCA)、判别分析(discriminant analysis, DA)、偏最小二乘法-判别分析(partial least squares discriminant analysis, PLS-DA)、人工神经网络(artificial neural networks, ANNs)等^[6-7]。本文就近年来化学模式识别方法在中药质量控制方面的应用展开综述, 为中药质量控制研究提供参考。

1 化学模式识别

1.1 无监督模式识别

1.1.1 CA CA属于无监督模式识别方法, 将样本数据在没有先验知识的前提下, 基于样本所表现的变量特征, 按照相似度进行归类。根据聚类途径不同, CA可分为系统聚类法(HCA)和动态聚类法^[8]。常用的CA方法包括HCA、K-均值聚类分析和

Ward's method等, 其中K-均值聚类分析属于典型的划分聚类分析方法, 其首先给予数据集初始划分, 再以此为起始点, 通过迭代不断调整样本归属, 最终达到最优目标函数, 进而获取聚类结果。Ward's method属于分层聚类分析方法的一种, 其着眼于将聚类分析作为1个方差分析的问题, 而不是通过测量距离或者关联度, 这种方法不需要计算集群中心, 而是最小化集群方差(相对于每个集群的质心)^[9-10]。CA常用于数据的初步探索性分析, 具有直观、结论形式简明的优点^[11]。然而聚类之前须对变量进行预处理, 采用不同的标准化方法获取的聚类结果不同。当样本量较大时, 聚类结论存在一定的困难, 无法对分类结果的准确性进行预测^[3, 11]。

1.1.2 PCA PCA^[12]为双线性模型方法, 利用方差最大原则, 对原始数据所包含的多个自变量进行线性拟合, 以新的低维变量代替原始高维变量, 即主成分, 各主成分之间互不相关, 从而这些主成分能够反映原始变量的绝大部分信息, 且所含的信息互不重叠, 进而实现数据的降维。高维数据可以通过以下方程进行降维。

$$X=CS^T+E$$

X 为原始数据, C 为得分矩阵, S^T 为载荷矩阵转置阵, E 为误差矩阵

每个样本对象在每个主成分上都有1个值, 同时每个变量在每个主成分上均有1个载荷值, 因此PC1与PC2的双标图(biplot)描述了最大的数据差异, 并提供了样本与荷载变量的比较, 可用于获取样本聚类的相似性以及荷载变量间、荷载变量与样本之间的关联等相关信息^[13]。PCA属于探索性数据分析方法(exploratory data analysis)的范畴^[14], 其采用少数综合变量来代替原始的众多变量, 具有可信度高、灵活性强, 侧重于综合评价信息的贡献影响力的特点^[3, 11]。但是当分组过多时, PCA可视化的分类效果则会大大降低; 原始数据间的相关性较弱时, 无法进行良好的降维, 获取的主成分无法综合地反映原始变量信息; 当存在许多未被检测的变量时, 如代谢组学研究中代谢产物只能在某些样本中发现, PCA获取的分析结果也会存在问题^[15]。然而PCA仍然是一种良好的探索性数据分析方法, 可为有监督的模式识别方法奠定分类依据。近年来, 该方法在中药质量控制方面具有广泛的应用。

1.2 有监督模式识别

1.2.1 SIMCA SIMCA^[16]是一种有监督的数据分

类方法,每个类别独立使用 PCA 建立模型,保留足够数量的主成分,进而产生了可用于每个类别的专属模型。因此,一类空间构建的边界可用于判别拟合分类模型样本与在选定置信区间内不属于该类别的样本。每个预测样本的信息可被划分成 2 个分量,一部分通过分类模型解释,另一部分通过残差解释,以表明预测样本与模型之间的拟合情况。1 个 SIMCA 模型由 1 个 PCA 模型集组成,数据集中每个分类均有 1 个 PCA 模型,每类可具有不同的主成分数目,主成分数目取决于相应类别的数据。每个 PCA 子模型包含了所有 PCA 通常部分,如均值向量、缩放信息、预处理(平滑、衍生等)。SIMCA 是建立在 PCA 基础上的有监督模式识别方法,原始数据的变量数和样本数的比例并非十分关键,在计算过程中 PCA 对其结果起到决定性作用。复杂的数据和不确定的测量信息均会导致显著主成分数目的估计困难,对于每个目标分类,主成分过多或者过少均会降低 SIMCA 模型的分类及预测能力,该缺点在一定程度上影响了 SIMCA 的广泛使用^[17-18]。

1.2.2 PLS-DA PLS 是一种寻找独立变量 X 与独立变量 Y 之间基本联系的方法,这个模型不仅考虑了 X 变量,同时也考虑了 X 变量与 Y 变量之间的关联,以 PCA 为数学基础,能够在自变量存在多重相关性的条件下进行回归建模。PLS-DA 是 PLS 的一种转化形式,采用分类响应变量 Y 来提高类别间的分离,计算校正模型的数据包括校正误差均方根、交叉验证误差均方根及测定的相关系数^[13]。PLS-DA 是一种有监督的模式识别方法,适用于区分两类及更多类别的样本,同时具备线性判别分析的分类能力又兼具 PLS 降维、降噪的优势^[19]。此外,PLS-DA 的优势还在于其更强大的解释能力,尤其是在原始变量方面。

1.2.3 支持向量机 (support vector machine, SVM) SVW 是一种通用的,特别是用于非线性数据的分析方法,基于使用 1 个核函数将原始数据映射到 1 个高维特征空间,并在该空间构建最佳的分类超平面^[20]。SVM 主要思想是针对两种分类问题,目标是在高维空间寻找 1 个最优分类超平面,以保证最小的分类错误率。SVM 建模的目的是确定在 2 个分类之间具有最大余量的最佳超平面,其可以应用于线性和非线性数据集和分离多个类。支持向量机判别分析 (SVMDA) 是建立在 SVM 基础上的一种有监督的分类方法,通过在 2 个类之间生成超平面边

界,将样本表示为空间中的点,超平面使类之间的距离最大化,并且与校准集中的样本的概率分布无关。SVMDA 包括 v -支持向量分类 (v -support vector classification) 和 c -支持向量分类 (c -support vector classification) 2 种分类方法,可较好地解决小样本、高维数、非线性和局部极小点等问题^[21]。然而,这种方法通过适当的核函数将原始数据指向了更高维的空间^[19]。

1.2.4 ANNs ANNs 也称为神经网络,是 1 个有向拓扑结构和学习规则的动态处理系统,具有自学习、自适应、联想存储、高速寻找优化解等功能优势^[22]。目前应用最普遍的 ANNs 是误差反向传播人工神经网络 (back propagation artificial neural network, BP-ANN)、自组织映射神经网络 (self organising maps, SOMs, 也称 Kohonen 网络) 和对偶传播人工神经网络 (counter-propagation artificial neural networks, CP-ANNs)。BP-ANN 是典型的多层网络,分为输入层、隐含层和输出层,层与层之间多采用全互连方式,同一层单元间不存在相互连接^[23]。Kohonen 网络是自组织竞争神经网络的一种,该网络为无监督性学习网络,能够识别环境特征并自动聚类;CP-ANNs 是基于 Kohonen 网络发展起来的,同时能够处理非监督性和监督性分类问题^[7]。基因算法 (遗传算法) 是一种自适应启发式群体型、概率性迭代全局收敛算法,可优化神经网络中神经元与迭代次数,使神经网络能够快速收敛,并且可以避免过训练问题^[24-25]。此外,径向基函数 (radial basis function, RBF) 神经网络,在中药质量、指纹图谱研究及药材真伪优劣鉴别中也得到了广泛的应用^[26]。

2 应用软件

SPSS 统计软件是目前最常用的统计分析软件之一,可实现 PCA、CA、Ward's method、RBF 神经网络、因子分析等多种模式识别分析^[26-27]。矩阵实验室 (Matlab) 具有语言简洁紧凑、运算符丰富、程序设计自由度大、程序可移植性好、数据可视化简单、拥有强大的工具箱等特点,是一种实用性强的科学计算工具,通过程序的编写及工具箱可实现大量的化学模式识别分析,如 PCA、BP-ANN、Kohonen 网络、CP-ANNs、SVM、PLS-DA、SIMCA、线性判别分析 (LDA)、K-最近邻法 (KNN) 等,其中 Kohonen 网络、CP-ANNs 可通过 Kohonen and CP-ANN toolbox 实现 (<http://micchem.disat.unimib.it/chm/download/download.htm>)^[13,28-29]。多元数据

分析软件 SIMCA 是 1 款过程分析和质量设计的工具软件,能够从众多数据中获取关键信息,可实现 PCA、CA、PLS-DA、正交偏最小二乘法判别分析 (OPLS-DA)、SIMCA 等多种模式识别分析,该软件无需程序编写,具有操作简单、方便快捷的特点^[30-31]。

3 化学模式识别在中药质量控制中的应用

3.1 产地来源

我国疆域辽阔,具有复杂的自然地理环境,日照、温度、土质等环境不尽相同,为药用植物及动物的生长提供了必要的有利条件,同时也使中药材的生产和质量具有一定的地域性^[32]。这也是古今医家都非常重视“道地药材”的缘故。如表 1 所示,近年来化学模式识别方法广泛应用在中药不同产地的研究中,常联合的仪器分析方法包括 HPLC、核磁共振光谱 (NMR)、红外光谱 (IR)、直接电离质

谱 (DI-MS)、电感耦合等离子体质谱法 (ICP-MS) 等,化学模式识别方法涉及 CA、PCA、DA、KNN、PLS-DA、最小二乘支持向量机 (LS-SVM) 等,其中最常用的化学模式识别方法为 PCA。

3.2 基原

中药基原是影响中药质量与疗效的最基本的内在因素之一。中药基原物种混乱直接影响到临床用药的有效性和安全性,对其进行基原鉴别,从源头上控制质量,是中药资源开发、中药生产及中药质量标准制定的必要环节^[43]。葛根与粉葛长期以来均作为葛根药材使用,两者基原不同,分别来源于豆科植物野葛 *Pueraria lobata* (Willd.) Ohwi 和甘葛藤 *Pueraria thomsonii* Benth. 的干燥根,自《中国药典》2005 年版开始将二者单独列出^[44]。Wong 等^[19]采用高效薄层色谱法 (HPTLC) 与 UPLC 对葛根和粉葛

表 1 化学模式识别方法在中药不同产地研究中的应用

Table 1 Application of chemical pattern recognition in CMM from different geographical origins

名称	产地数	分析成分	分析方法	模式识别	软件	文献
浮萍 <i>Spirodela polyrriza</i> (L.) Schleid.	8	5 个黄酮类成分	HPLC	HCA+PCA	SPSS	33
丹参 <i>Salvia miltiorrhiza</i> Bge.	5	13 个共有峰	UPLC-Q-TOF/MS	HCA+PCA	SPSS	34
天麻 <i>Gastrodia elata</i> Bl.	10	DI-MS 中归一化强度大于 5% 的峰	DI-MS	PCA	Simca	35
甘草 <i>Glycyrrhiza uralensis</i> Fisch.、 胀果甘草 <i>Glycyrrhiza inflata</i> Bat.、 光果甘草 <i>Glycyrrhiza glabra</i> L.	4	5 个甘草皂苷类成分	HPLC	PCA	SPSS	36
附子 <i>Aconitum carmichaelii</i> Debx.	4	409 个质谱峰	UPLC-Q-TOF/MS	PCA+PLS-DA	Simca-P	30
独一味 <i>Lamiophlomis rotata</i> (Benth.) Kudo	3	19 个化学成分	NMR 光谱	HCA+PCA	Simca-P	31
何首乌 <i>Polygonum multiflorum</i> Thunb.	6	22 个共有峰	UPLC-Q-TOF/MS	PCA+Ward's method+CP-ANN	Matlab	29
丹皮 <i>Paeonia suffruticosa</i> Andr.	3	14 个化学成分, 7 个金属元素	HPLC-MS、ICP-MS	PCA+LDA+KNN+BP-ANN+LS-SV	Matlab	16
白花丹参 <i>Salvia miltiorrhiza</i> Bge. var. <i>alba</i> C. Y. Wu et H. W. L.	4	4 000~10 000 cm ⁻¹ 光谱区域	近红外光谱	PCA+DA	Matlab	37
苍耳子 <i>Xanthium sibiricum</i> Patr.	10	9 个共有峰	IR 指纹图谱	HCA	SPSS	38
杜仲 <i>Eucommia ulmoides</i> Oliv.	22	901 个质谱离子	直接进样-质谱检测 指纹图谱	PCA+PLS-DA	Simca-P	39
金银花 <i>Lonicera japonica</i> Thunb.	6	17 个微量元素	ICP-MS	SOM	Matlab	40
滇重楼 <i>Paris polyphylla</i> var. <i>yunnanensis</i> (Franch.) Hand. -Mazz.	19	8 种皂苷成分	HPLC	HCA+PCA	SPSS	41
车前子 <i>Plantago asiatica</i> L.	24	10 个共有峰	HPLC	HCA	SPSS	42

进行分析,利用 KNN、PLS-DA、PCA-DA、SVM-DA、SIMCA 5 种化学模式识别方法比较 2 种色谱法的分类效果,结果显示 HPTLC 与 UPLC 分类模型结果具有相似性,前 4 种模式识别方法对于葛根和粉葛具有良好的识别率,而 SIMCA 模型的分类效果最差。苏湮等^[45]采用 HPLC 结合 PCA,对贡菊、滁菊、杭菊和亳菊 4 种基原的药用菊花中特征成分进行分析,该方法能大体上将不同来源的菊花按基原进行分类,为菊花的基原鉴定提供了研究思路。

3.3 炮制

中药炮制蕴含着中医整体观和辨证论治理念,是中医药学的特色之一,也是中药临床应用的特点和必然要求,中药通过炮制可达到增效减毒、改变药物性能、引药入经及扩大应用范围等目的^[46]。Zhou 等^[47]采用 HPLC 指纹图谱结合 HCA、PCA、PLS-DA 3 种化学模式识别方法对炮制前后的播娘蒿种子进行了比较分析,研究表明生品与炮制品的化学成分具有一定的差异性,炮制后产生了许多新的化学成分,并获取了重要差异成分,进而阐明了播娘蒿的炮制机制。孙立丽等^[48]建立了何首乌 UPLC 指纹图谱,并在此基础上成功建立了 5 种化学模式识别(PCA、CP-ANN、PLS-DA、SVMDA、CP-ANN)模型,并采用有监督的 CP-ANN 模式识别模型下获取的 Kohonen 权重值为研究指标,探讨了何首乌炮制前后的整体化学成分变化,为何首乌的炮制研究提供了研究思路与实验方法。

3.4 真伪鉴别

中药真伪问题是影响中药质量的另一重要因素,“真”即正品,指国家药品标准所收载的中药品种,“伪”即伪品,凡不符合国家药品标准规定的中药品种均应视为伪品^[49]。女贞子来源于木犀科植物女贞 *Ligustrum lucidum* Ait. 的干燥成熟果实,收载于《中国药典》2015 年版一部中,为养阴常用中药之一。日本女贞 *Ligustrum japonicum* Thunb. 与女贞来源于同科植物,果实性状极为相似,但日本女贞误用后可引起中毒,曹运姣等^[50]建立了紫外谱线组法-化学模式识别鉴别女贞子与日本女贞的方法,在 PCA 的基础上对样本进行 Q 型聚类分析,结果表明二者之间有显著性差异,且分类界限明显;建立了 Fisher 判别分析,可用于女贞子与日本女贞子的鉴别。

3.5 其他应用

Shan 等^[51]采用化学模式识别方法综合评价 2~6 年生于不同季节采收的人参 HPLC 指纹图谱,建

立的 PCA、HCA 2 种化学模式识别方法为不同生长年限及采收季节的人参样本提供了质量比较和有效的区分方法。顾志荣等^[52]利用 ¹³C-NMR 特征图谱结合 PCA、模糊聚类分析(FCA)建立不同生长年限当归的鉴别与质量评价方法,能够对来源于不同生长年限的当归进行快速、准确地鉴别。汤彦丰等^[53]利用 BP-ANN 对傅里叶红外光谱法数据进行处理,该方法可以有效地鉴别野生和栽培的紫花地丁。

3.6 化学模式识别应用过程中的技术问题

3.6.1 数据预处理方法 在建立化学模式识别分类模型之前,通常要对原始数据进行数据预处理,使数据更有利于进行分类运算。标度化方法可使所有变量的变化幅度处于同一个水平上,常用的标度化方法包括范围标度化、自标度化、标准化、变换法、组合法等,其中标准化预处理方法又包括面积归一化、最大归一化等^[54]。不同的情况下可采用不同的数据预处理方法,采用的方法取决于数据的来源、问题的类型、化学或物理因素等,不可统而论之^[55]。基于 Matlab 的化学模式识别工具箱中,常用的数据预处理方法有范围标度化、自标度化、中心化等。PLS_Toolbox 具有强大的数据预处理功能,除常用的标度化方法外,还包括平滑、衍生、去趋势化、基线(加权最小二乘法)等预处理方法来解决噪音、偏移和基线问题;数据选择类型包括绝对值、log10;多种数据过滤方法包括正交信号校正方法、广义最小二乘加权等^[56]。

3.6.2 其他 数量充足、来源可靠、具代表性的中药样本是建立具有广泛应用价值、操作简便规范、优良的识别与预测能力化学模式识别模型的前提条件。建立化学模式识别模型所选的样本量依赖于所选模型、研究的主体、待解决的中药问题等多种因素。同时,由于化学模式识别方法各有其优缺点,若仅使用单一方法则难以避免局限性,因此对于中药质量控制研究,需要建立多种模式识别方法加以互补以达到最佳的质量控制结果。

4 结语与展望

中药本身化学成分复杂,基原品种、产地来源、生长环境、采收季节、加工炮制工艺等诸多因素进一步增加了中药质量控制与评价研究的难度。由于中药本身的特殊性,仅靠传统外观鉴别、显微鉴别及少数有效成分分析鉴定中药的真假优劣具有一定的不准确性^[57]。化学模式识别技术可对 HPLC、UPLC-Q-TOF/MS、IR、NMR 等多种现代仪器分析

获取的数据进行客观分析,既可对多个指标进行统计分析,又可将整个图谱信息数量化,进而使其可以被计算机识别与处理,从而可以更加客观地反映中药的质量信息,达到全面控制中药质量的目的。化学模式识别技术已在中药质量控制研究中得到了广泛应用,然而中药的化学模式识别研究工作依然存在很多问题,获取的中药数据量不够或样本信息不可靠是最常见的问题之一,但是毋庸置疑的是化学模式识别技术是一种有效评价中药质量的方法。相关中药样品库和数据库的建立,如中药多基原药材的鉴别方法数据库,以及来源信息确切的中药购置渠道,将大大推进化学模式识别技术在中药质量控制研究工作的顺利展开,对中药的质量控制与评价具有重要现实意义。

参考文献

- [1] 笪舫芳,朱 华,刘 颖. 中药现代化之路: 向左走,向右转 [J]. 中国医药导报, 2012, 9(16): 122-123.
- [2] 刘昌孝,陈士林,肖小河,等. 中药质量标志物(Q-Marker): 中药产品质量控制的新概念 [J]. 中草药, 2016, 47(9): 1443-1457.
- [3] 刘昌孝. 基于中药质量标志物的中药质量追溯系统建设 [J]. 中草药, 2017, 48(18): 3669-3676.
- [4] 马诗瑜,沈 岚,洪燕龙,等. 化学计量学在中药定量表征中的应用 [J]. 世界科学技术—中医药现代化, 2014, 16(12): 2700-2707.
- [5] 吴功煌,史新元,乔延江. 近红外模式识别技术在中药质量控制中的应用研究进展 [J]. 世界科学技术—中医药现代化, 2010, 12(2): 265-270.
- [6] 汪小莉,李 嫒,秦昆明,等. 近红外光谱学与化学计量学在中成药液体制剂过程分析中的应用 [J]. 中草药, 2013, 44(15): 2165-2171.
- [7] Ballabio D, Consonni V, Todeschini R. The Kohonen and CP-ANN toolbox: A collection of MATLAB modules for self organizing maps and counterpropagation artificial neural networks [J]. *Chemom Intell Lab Syst*, 2009, 98(2): 115-122.
- [8] 邱新华,唐铁鑫,刘 燕,等. 几种模式识别方法用于中药挥发油红外光谱法鉴别的比较研究 [J]. 中国药房, 2015, 26(21): 2986-2988.
- [9] 汪 涛,鲍远程. 聚类分析在中药复方研究中的应用 [J]. 成都中医药大学学报, 2013, 36(2): 121-122.
- [10] Lee A, Willcox B. Minkowski generalizations of ward's method in hierarchical clustering [J]. *J Classif*, 2014, 31(2): 194-218.
- [11] 李新蕊. 主成分分析、因子分析、聚类分析的比较与应用 [J]. 山东教育学院学报, 2007(6): 23-26.
- [12] 庞晶瑶,王伽伯,马致洁,等. 基于化学指纹图谱和生物毒性检测的何首乌质量评控 [J]. 中草药, 2014, 45(23): 3392-3396.
- [13] Ni Y, Mei M, Kokot S. One-and two-dimensional gas chromatography-mass spectrometry and high performance liquid chromatography-diode-array detector fingerprints of complex substances: A comparison of classification performance of similar, complex *Rhizoma Curcumae* samples with the aid of chemometrics [J]. *Anal Chim Acta*, 2012, 712(2): 37-44.
- [14] Gad H A, El-Ahmady S H, Abou-Shoer M I, et al. Application of chemometrics in authentication of herbal medicines: A review [J]. *Phytochem Anal*, 2013, 24(1): 1-24.
- [15] Brereton R G. *Chemometrics for Pattern Recognition* [M]. Chichester: John Wiley & Sons, Ltd., 2009.
- [16] Zhuang H, Ni Y, Kokot S. Combining HPLC-DAD and ICP-MS data for improved analysis of complex samples: Classification of the root samples from *Cortex moutan* [J]. *Chemom Intell Lab Syst*, 2014, 135(11): 183-191.
- [17] 龙草芳. SIMCA 分类法中主成分分析算法的研究 [J]. 数字技术与应用, 2017(4): 145-146.
- [18] 徐 路,付海燕,姜 宁,等. 基于偏最小二乘回归的类模型方法用于中药牛黄的真伪鉴别 [J]. 分析化学, 2010, 38(2): 175-180.
- [19] Wong K H, Razmovski-Naumovski V, Li K M, et al. Differentiating *Puerariae Lobatae Radix* and *Puerariae Thomsonii Radix* using HPTLC coupled with multivariate classification analyses [J]. *J Pharm Biomed Anal*, 2014, 95(3): 11-19.
- [20] Li H, Liang Y, Xu Q. Support vector machines and its applications in chemistry [J]. *Chemom Intell Lab Syst*, 2009, 95(2): 188-198.
- [21] 林兆洲,史新元,郭明星,等. 清开灵注射剂清热相关非线性生物标记物的辨识及可视化方法研究 [J]. 中华中医药杂志, 2014, 29(5): 1328-1333.
- [22] Ballabio D, Vasighi M. A MATLAB toolbox for self organizing maps and supervised neural network learning strategies [J]. *Chemom Intell Lab Syst*, 2012, 118(2): 24-32.
- [23] 覃洁萍,王智猛,李梦龙. 化学计量学方法在中药鉴别及质量控制方面的应用 [J]. 数理医药学杂志, 2004, 17(4): 351-355.
- [24] Ballabio D, Vasighi M, Consonni V, et al. Genetic algorithms for architecture optimisation of counter-propagation artificial neural networks [J]. *Chemom Intell Lab Syst*, 2011, 105(1): 56-64.
- [25] Hakimzadeh N, Parastar H, Fattahi M. Combination of multivariate curve resolution and multivariate classification techniques for comprehensive high-performance liquid chromatography-diode array absorbance detection fingerprints analysis of *Salvia reuterana* extracts [J]. *J Chromatogr A*, 2014, 1326: 63-72.
- [26] 邹慧琴,李 硕,陶 欧,等. RBF 神经网络在中医药领域中的应用及其在 SPSS 17.0 软件中的实现 [J]. 中华中医药学刊, 2015, 33(2): 336-338.

- [27] 韦卓纯, 姚志红, 王其意, 等. UPLC 结合化学计量学方法的肿节风指纹图谱研究 [J]. 中草药, 2015, 46(6): 895-900.
- [28] 段慧琦, 陆晓华. MATLAB 在化学计量学中的应用 [J]. 分析科学学报, 2003, 19(1): 86-91.
- [29] Sun L L, Wang M, Zhang H J, *et al*. Comprehensive analysis of *Polygoni Multiflora Radix* of different geographical origins using ultra-high-performance liquid chromatography fingerprints and multivariate chemometric methods [J]. *J Food Drug Anal*, 2016, doi: 10.1016/j.fda.2016.11.009.
- [30] 张定堃, 韩雪, 李瑞煜, 等. UPLC-Q-TOF-MS 分析不同产地泥附子化学成分的差异 [J]. 中国中药杂志, 2016, 41(3): 463-469.
- [31] Pan Z, Fan G, Yang R P, *et al*. Discriminating *Lamiophlomis rotata* according to geographical origin by ¹H-NMR spectroscopy and multivariate analysis [J]. *Phytochem Anal*, 2015, 26(4): 247-252.
- [32] 张乐生. 浅谈中药的产地与采集 [J]. 内蒙古中医药, 2011, 29(9): 77.
- [33] 蒋祥亮, 刘聪燕, 陈彦, 等. 不同产地浮萍中 5 种黄酮成分及抗氧化活性的化学模式识别研究 [J]. 中草药, 2017, 48(5): 985-990.
- [34] 曾慧婷, 沙秀秀, 宿树兰, 等. 不同产地丹参茎叶 UPLC 指纹图谱与化学模式识别研究 [J]. 中草药, 2017, 48(4): 767-772.
- [35] Wong H Y, Hu B, So P K, *et al*. Rapid authentication of *Gastrodiae Rhizoma* by direct ionization mass spectrometry [J]. *Anal Chim Acta*, 2016, 938: 90-97.
- [36] 方诗琦, 瞿其扬, 仲欢欢, 等. 甘草中 5 种三萜皂苷的同时测定及主成分分析 [J]. 中成药, 2016, 38(2): 336-341.
- [37] Duan X J, Zhang D L, Nie L, *et al*. Rapid discrimination of geographical origin and evaluation of antioxidant activity of *Salvia miltiorrhiza* var. *alba* by fourier transform near infrared spectroscopy [J]. *Spectrochim Acta A Mol Biomol Spectrosc*, 2014, 122(6): 751-757.
- [38] 徐海星, 陈慧, 许沛虎, 等. 不同产地苍耳子红外指纹图谱鉴别 [J]. 中药材, 2014, 37(6): 971-973.
- [39] 赵杨, 周欣, 龚小见, 等. 直接进样-质谱检测指纹图谱技术结合化学模式识别法研究杜仲的质量 [J]. 中草药, 2014, 45(19): 2844-2849.
- [40] 申明金. 基于自组织特征映射神经网络的金银花分类研究 [J]. 化学分析计量, 2013, 22(2): 35-37.
- [41] 刘欢, 何忠俊, 梁社往, 等. 不同产地滇重楼 HPLC 指纹图谱及化学模式识别研究 [J]. 云南农业大学学报: 自然科学, 2013, 28(1): 88-95.
- [42] 曾金祥, 罗光明, 朱玉野, 等. 车前子化学模式识别及芦丁含量测定研究 [J]. 中药材, 2010, 33(8): 1237-1239.
- [43] 李西林, 陈科力, 李薇, 等. 试析中药鉴定中的基原鉴定问题 [J]. 中药材, 2009, 32(12): 1929-1932.
- [44] 李生茂, 刘琳, 谭睿, 等. 葛根和粉葛高效液相色谱指纹图谱及化学模式识别 [J]. 辽宁中医杂志, 2016, 43(5): 1034-1037.
- [45] 苏涔, 王京辉, 郭洪祝, 等. 中药菊花基原特征成分的分析 [J]. 药物分析杂志, 2016, 36(1): 59-67.
- [46] 张夏维, 何家振, 周贤, 等. 浅谈中药炮制中的整体观念与辨证论治观 [J]. 时珍国医国药, 2017, 28(3): 671-672.
- [47] Zhou X D, Tang L Y, Wu H W, *et al*. Chemometric analyses for the characterization of raw and processed seeds of *Descurainia sophia* (L.) based on HPLC fingerprints [J]. *J Pharm Biomed Anal*, 2015, 111: 1-6.
- [48] 孙立丽, 任晓亮, 张慧杰, 等. 基于超高效液相色谱指纹图谱和化学计量学相结合的何首乌炮制研究 [J]. 中华中医药杂志, 2017, 32(5): 2194-2197.
- [49] 向丽, 张贵君, 王晶娟. 中药真伪优劣鉴定的研究思路与方法 [A] // 第四届中国中药商品学术大会暨中药鉴定学科教学改革与教材建设研讨会论文集 [C]. 北京: 中国商品学会, 2015.
- [50] 曹运姣, 王旭, 容蓉. 女贞子与日本女贞子的 UV-化学模式识别 [J]. 中国实验方剂学杂志, 2016, 22(17): 32-36.
- [51] Shan S M, Luo J G, Huang F, *et al*. Chemical characteristics combined with bioactivity for comprehensive evaluation of *Panax ginseng* C. A. Meyer in different ages and seasons based on HPLC-DAD and chemometric methods [J]. *J Pharm Biomed Anal*, 2014, 89(4): 76-82.
- [52] 顾志荣, 杨应文, 王亚丽, 等. 基于主成分分析和模糊聚类分析的不同生长年限当归 ¹³C 核磁共振特征图谱 [J]. 中国医院药学杂志, 2014, 34(24): 2083-2087.
- [53] 汤彦丰, 王绍茹, 王志宝, 等. 紫花地丁红外光谱的人工神经网络鉴别研究 [J]. 湖北农业科学, 2013, 52(7): 1656-1658.
- [54] 梁逸曾, 俞汝勤. 化学计量学 [M]. 北京: 高等教育出版社, 2003.
- [55] 史永刚, 冯新沪, 李子存. 化学计量学 [M]. 北京: 中国石化出版社, 2002.
- [56] Wise B M, Gallagher N B, Bro R, *et al*. *Chemometrics Tutorial for PLS_Toolbox and Solo* [M]. Washington: Eigenvector Research, Inc., 2006.
- [57] 刘东方, 赵丽娜, 李银峰, 等. 中药指纹图谱技术的研究进展及应用 [J]. 中草药, 2016, 47(22): 4085-4094.