

基于数据挖掘的中成药不良反应监测研究进展及探讨

徐 枫¹, 王 晨¹, 李中良², 王光安^{3*}, 姚雅淇³

(1. 河南推拿职业学院, 河南 洛阳 471023; 2. 河南推拿职业学院附属医院, 河南 洛阳 471023; 3. 河南中医药大学针灸推拿学院, 河南 郑州 45008)

摘要: 中成药临床安全性是影响中药产业发展的核心问题, 开展中成药不良反应监测是保障临床安全应用的关键。中成药安全性监测研究需要不断吸收创新方法与理念。随着信息技术的发展与大数据时代的到来, 以真实世界医疗数据为基础的数据挖掘研究已成为中成药不良反应监测的趋势。本文总结了近年来采用数据挖掘方法开展的中成药不良反应监测研究, 阐述了不同方法的原理, 包括比值失衡法、处方序列分析方法、机器学习等, 并对这些算法在中成药不良反应监测研究中的应用情况进行总结。同时, 探讨了中成药不良反应监测数据挖掘存在的关键共性问题, 主要包括多中心医疗数据共享、引发药品不良反应的质量因素辨析、数据挖掘技术人才队伍建设等, 以期为其进一步发展提供参考。

关键词: 中成药; 药物警戒; 药品不良反应监测; 数据挖掘

中图分类号: R288

文献标志码: A

文章编号: 1001-1528(2024)12-4089-06

doi: 10.3969/j.issn.1001-1528.2024.12.029

安全性是药品最基本的属性, 也是医疗卫生和监管机构重点关注的问题。自沙利度胺事件后, 世界各国开始认识到药品安全性评价与药品不良反应监测的重要性。随着我国医疗卫生体制的发展及药物警戒制度的完善, 药品不良反应监测已经成为药品生产企业、医疗机构和药政部门安全责任体系中的重要组成部分。中医药是我国生物医药领域具有战略地位的支柱产业之一, 随着中成药临床使用数量的增长, 不良反应/不良事件报道逐渐增多。近三年发布的国家药品不良反应监测年度报告显示, 中药不良反应/事件数量呈现逐年上升趋势^[1-3]。因此, 有必要持续开展中成药不良反应监测研究, 为临床安全用药提供指导。近年来, 科研机构在中成药安全性评价及不良反应监测技术研究方面做了大量工作^[4]。药品不良反应监测是涉及药学、临床医学、统计学等多学科交叉的系统工程。随着新一代信息技术的飞速发展, 以药品上市后医疗数据为载体开展数据挖掘研究, 已经成为中成药不良反应监测的发展方向。本文基于中成药安全性评价现状, 分析中成药不良反应监测对数据挖掘的需求; 对借助数据挖掘技术开展中成药不良反应监测研究的进展进行综述, 探讨中成药开展不良反应数据挖掘存在的关键共性问题, 以期为其进一步发展提供参考。

1 利用数据挖掘开展中成药不良反应监测的必要性

1.1 中成药开展不良反应研究的迫切性 中成药的研制大多是在我国传统中医药理论指导下, 沿袭古代经典名方和经验方, 或在此基础上改良而来。然而, 传统中医药理论与中药典籍缺少对处方制剂潜在安全性风险的描述与记载。

虽然在上市前会开展非临床毒性评价研究, 但是中药化学成分复杂, 其中引起不良反应的物质基础及毒性机理难以被充分解释。由于缺乏这方面的研究, 中成药说明书中的“不良反应”项往往显示“尚不明确”, 给临床合理用药造成较大困难。2009 年原国家食品药品监督管理局发布《关于开展中药注射剂安全性再评价工作的通知》, 任务之一是加强中药注射剂不良反应监测和分析、反馈工作, 指导企业修订说明书, 促进临床合理用药。2023 年 2 月, 国家药品监督管理局发布了《中药注册管理专门规定》, 其中明确指出, 中药说明书“禁忌”“不良反应”“注意事项”中任何一项在本规定施行之日起满 3 年后申请药品再注册时仍为“尚不明确”的, 依法不予再注册^[5], 这就从法规层面要求生产企业必须开展中成药不良反应的研究, 完善中成药不良反应相关信息, 从而指导临床合理用药。

1.2 中成药不良反应监测模式 药品上市前开展的临床试验在安全性评价方面, 由于存在伦理学限制以及样本代表性不足等问题, 导致难以反映实际药品不良反应发生情况, 尤其是罕见或长期暴露后迟发的不良反应往往无法及时发现^[6-7]。因此, 药品不良反应监测更多依靠上市后临床安全性评价研究。目前, 我国中成药上市后不良反应主要依赖自发呈报系统被动监测, 但是自发呈报存在迟报或漏报、报告质量不满足要求、可利用性不足等问题^[8-9]。前瞻性医院集中监测被认为是中成药上市后临床安全性评价最成熟的方法^[4]。中国中医科学院团队采用医院集中监测方法开展了多个中成药品种的不良反应研究^[10-13], 形成了行业技术规范^[14]。然而, 医院集中监测需要严格的设计方案, 并

收稿日期: 2024-02-02

基金项目: 河南中医药大学特色骨干学科项目 (STG-ZYX02-202114)

作者简介: 徐 枫 (1981—), 男, 讲师, 研究方向为中医学。E-mail: xufengtc@163.com

* 通信作者: 王光安 (1978—), 男, 博士, 副教授, 研究方向为针灸治疗的疗效评价。E-mail: wangganpharm@163.com

且有专门的人员对患者在医院用药期间进行全程监测^[15]，导致该方法需要耗费大量的研究经费和人力，研究过程质量控制难度较高。

随着新一代信息技术及“物联网+医疗服务”模式的发展，来自真实世界的医疗健康大数据快速形成，以其为载体开展上市后安全性评价研究为药品不良反应监测带来了新的机遇^[16]，欧美等国家及地区已经基于药品上市后医疗数据资源建立了安全性主动监测系统^[17]。随着相关政策的颁布实施、医疗数据的不断积累和新技术的涌现，我国已经具备药品上市后开展安全性监测研究的条件^[18]。真实世界的电子医疗数据具有数据量大、来源广泛、时效性强等优势，直接利用已有数据能够节约研究成本，提高研究效率，在中成药不良反应监测中有良好的应用前景。

1.3 中成药不良反应监测对数据挖掘技术的需求 真实世界医疗数据能够提供丰富的信息，但是这些数据所含变量类型众多，结构化、半结构化与非结构化数据混杂，采用常规统计分析技术难以对其处理。中成药的原料来自天然产物，化学成分众多，且受到中药制造工艺水平的影响，导致不同批次间质量不够稳定^[19]；而中成药，特别是中药注射剂，在临床上往往与其他药品联合使用^[20]。这些问题使得中成药临床不良反应的成因模糊不清，从而加剧了中成药医疗数据的复杂程度。这就对数据分析技术水平提出了更高的要求，有必要引进创新思维与方法开展中药医疗大数据的研究。

数据挖掘是利用计算机技术和信息技术的发现、推理和思维过程，具有自动预测趋势和行为、关联分析、聚类、类与概念描述、偏差检测等功能，从而实现将数据提升为有价值的知识^[21]。数据挖掘能够处理结构复杂、关系模糊的大数据集，发现数据背后隐藏的模式与规律，已经在药品不良反应监测研究中得到应用^[22]。在中药领域，数据挖掘可用于中药处方配伍规律研究^[23]，也可对中成药生产过程质量数据进行回顾分析^[24]，说明数据挖掘在针对复杂中药体系问题的研究中有良好的应用潜力。因此，以医疗大

数据为载体，借助数据挖掘开展上市后安全性评价，有望成为中成药不良反应监测研究的技术趋势。

2 数据挖掘在中成药不良反应监测中的应用

随着大数据技术的兴起与推广，数据挖掘技术在中成药上市后不良反应监测研究中逐渐得到重视与的应用。研究采用的信息收集方式包括来自真实世界医疗数据库的回顾性分析，也包括前瞻性的医院集中监测以及自发呈报系统数据库。

2.1 比例失衡方法 比值失衡法是较早的一种不良反应信号挖掘方法，多用于分析自发呈报系统数据库。该方法通过估计不良反应报告系统中实际出现的与某种药物有关的不良反应数量与预期数量或者与其他药物导致的其他不良反应数量的比值来预估不良反应发生情况，当比值大到一定程度（失衡）时，认为有可能是不良反应信号^[25]。在实际使用时，一般需要将资料整理成四格列联表形式，当某个药品目标不良反应数量的比值统计结果超出临界值时，表明生成不良反应信号。比值失衡方法主要包括采用频数分析的报告优势比法、比例报告比值比法、贝叶斯置信区间递进神经网络法等。夏旭东等^[26]采用报告优势比法、综合标准法（将比例报告比值比法与卡方检验结合的方法）、贝叶斯置信区间递进神经网络法对清开灵注射液不良反应进行研究，最终挖掘出 11 个新的不良反应信号。在一项对双黄连注射液不良反应信号监测研究中，研究人员采用报告优势比法、比例报告比值比法和贝叶斯置信区间递进神经网络法对自发报告系统中的数据进行了评价。结果显示，报告优势比法和比例报告比值比法对不良反应预警一致性较高。当某种不良反应报告数较多时，3 种方法的预警结果一致；但是，当报告数量较少时，报告优势比法和比例报告比值比法估计误差大，容易出现假阳性，贝叶斯置信区间递进神经网络法能够很好地克服这一问题^[27]。报告优势比法和比例报告比值比法属于传统的频数计算方法，而贝叶斯置信区间递进神经网络法运用贝叶斯原理建立，与频数法存在一定的差异^[27-28]，2 种方法优缺点见表 1。

表 1 频数法与贝叶斯置信区间递进神经网络法的优势和局限性

方法	优势	局限
报告优势比法和比例报告比值比法	原理简单,计算简便,易于理解,分析结果灵敏度高; 当药物涉及的不良反应报告较多时,该方法适用	特异度低,信号结果不稳定,容易受个别值影响; 不良反应报告数量较少时,预警信号偏差大,容易出现假阳性
贝叶斯置信区间递进神经网络法	分析结果特异性高,信号结果稳定; 有利于动态观察不良反应的变化,适用于探索预警信号的时间趋势	计算过程复杂,分析结果灵敏度一般

2.2 处方序列分析方法 处方序列分析法是一种基于药品处方数据集来评价药品临床安全性的方法。当某个中成药出现不良反应，并且需要其他药物来针对该不良反应进行治疗时，可以采用该方法。其基本原理是当患者服用某种特定药物（指示药）后出现某种不良反应，需要另外一种特定药物（标签药）针对该不良反应进行治疗，指示药与标签药在一定时间内会呈现明显的先后顺序模式，通过分

析处方数据库中指示药与标签药在这段时间内所表现出的频率分布特征，来评价药物与不良反应之间是否存在关联^[29]。在一项评价参芪扶正注射液临床安全性影响因素的研究中，研究人员采用回顾性设计收集来自 20 家医院信息系统中的处方数据，根据使用参芪扶正注射液后是否使用地塞米松注射液进行处方序列分析，判断是否发生疑似过敏反应并开展分组研究。通过统计检验，参芪扶正注射液

单次使用剂量较大或合并使用苦参碱注射液与疑似过敏反应显著相关，建议严格按照说明书剂量用药^[30]。

来自真实世界的医疗大数据没有经过随机化处理，数据内部存在偏倚。因此，在分析前需要经过设计来减小混杂与偏倚对结果的影响。吕建等^[4]提出，在医疗数据库和处方序列分析法基础上，将倾向性评分匹配法与巢氏病例对照设计相结合来处理混杂因素，从而提高统计效应的准确度，形成了一套基于真实世界的中成药安全性评价方法学体系。已经有报道采用上述方法完成了多项中药注射剂疑似过敏反应影响因素的研究^[31-32]，进一步明确了中药注射剂临床应用的风险因素。目前已经报道的中成药研究中，仅采用处方序列分析法判断是否发生不良反应并作为对研究样本进行分组依据，数据分析过程仍采用其他方法完成。

在处方序列分析法基础上发展而来的处方序列对称分析法可以通过评价指示药与标签药分布的对称性来判断目标药物是否与不良反应存在关联，是一种快速挖掘不良反应信号的方法^[33]。在具体应用时，先从数据库中采集一定时间范围内处方了指示药和标签药的患者信息。根据 2 种处方药的先后顺序对患者分组，将先处方指示药后处方标签药的人数，除以先处方标签药后处方指示药的人数，得到粗略的序列比。CSR 可以反映 2 组不对称程度，但是没有考虑处方随时间变化趋势。为了调整这种偏倚，先计算假定没有任何因果关联情况下期望的序列比，即空间效应序列比。最后，以粗略的序列比与空间效应序列比的比值得到调整序列比，当其统计学显著性下限大于 1 时，表明存在潜在的因果关系^[34]。相比于包含大量混杂因素的传统流行病学设计，处方序列对称分析法使用较少变量即可完成不良反应信号挖掘^[4]，适用于大型医疗保险数据库和药品处方数据库。在一项对喜炎平注射液风险性信号评价中，研究人员以全国基本医疗保险数据库为来源，对处方了喜炎平注射液与抗过敏药物的患者数据进行处方序列对称分析。结果显示，喜炎平注射液具有引起过敏反应的风险。本研究纳入了同一天处方指示药和标签药的患者，有助于强化安全性信号，并且仅需患者编号和 2 种药物的首次处方信息便可以完成初步分析^[35]。在药品不良反应信号发掘研究中，处方序列对称分析法具有原理简单、方法容易实现、对混杂因素比较稳健、能捕捉罕见不良反应信号等优势^[36]。该方法计算速度快，对数据集要求较低，能够迅速进行风险评估与测量，被认为适合中药注射剂不良反应又急又快特点^[36]。

2.3 基于机器学习的数据挖掘 机器学习是一门多领域交叉学科，其内容涉及计算机科学、统计学、概率论等多门学科，作为大数据时代与人工智能领域发展最快的技术之一，已经在不同场景复杂数据问题处理中取得诸多进展。机器学习的工程化概念为计算机程序利用经验来学习任务，如果针对任务的性能随着经验不断增长，则称为机器学习^[37]。该技术可以不依赖系统的基本原理和理论，从体量

庞大、变量众多、结构多样的数据中发现不太明显的规律，提取新的知识。机器学习对数据的处理是通过算法实现的，不同算法根据其原理不同，所解决的问题类型也有所区别。

Logistic 回归是一种二分类算法，被广泛用于估算一个对象属于某种类别的概率。在药品不良反应信号检测时，对是否发生特定不良反应的样本进行分组，以影响因素作为自变量对目标不良反应建立 Logistic 回归模型，计算模型各项的优势比，根据优势比值统计学显著性来估计影响因素的相对危险程度。一般可采用单因素回归或多因素回归的方法来建立模型。吉萌萌^[38]采用前瞻性医院集中监测方式开展了清开灵注射液不良反应影响因素研究。以巢氏病例对照设计划分不良反应组与对照组，分别采用单因素与多因素 Logistic 回归建立模型，结果显示肿瘤患病史、皮疹不良反应史、用药时患有胃炎疾病以及中风病是清开灵注射液发生药品不良反应的风险因素，而与中药联合用药发生不良反应的风险不明显。一项对 7 种活血化瘀类中药注射剂不良反应影响因素研究以回顾性设计收集住院患者临床资料，通过 Logistic 回归判断出临床用药的风险因素，并得出联合用药时冲管是保护性因素，根据该结果给出了合理用药建议^[39]。Logistic 回归能够定量评估与药品不良反应发生相关的风险因素，在多个中药注射剂该问题的研究中得到应用^[40-41]。此外，Logistic 回归也可与其他数据挖掘方法联合使用，例如先采用报告优势比法检测丹参川芎嗪注射液不良反应信号，在此基础上使用 Logistic 回归筛选出风险因素^[42]。

关联规则分析用来在大规模数据集中发现事件之间的潜在关系，关联规则和频繁项集是该方法的核心概念。其中，关联规则表示 2 种或多种事件之间存在较强的关系，频繁项集表示同时出现次数较多的集合^[43]。关联规则的强度需要通过支持度和置信度来评估。在药品不良反应信号研究中，支持度表示某种药品与某不良事件同时发生的概率，置信度表示使用某种药品后某不良事件出现的概率^[43]。Apriori 是经典的关联规则算法，基于数据库中联合用药变量建立关联规则模型，对数据集进行逐层搜索，以迭代的方式识别所有频繁项集形成关联规则^[20]。在中成药不良反应监测研究中，Apriori 可用于评估不良反应相关的影响因素，例如对热毒宁注射液不良反应发生类型、严重程度和结果进行关联分析^[44]。Apriori 也可用于研究不良反应的流行病学特征，例如评估双黄连注射液可能引发的多系统损害特征，并对与不良反应相关的患者状况以及临床用药因素进行分析^[45]。Apriori 算法原理架构简单、对数据要求低，能够量化表征临床变量间相互影响的复杂关系。但是，分析过程需要对数据库全部记录反复扫描，导致计算量庞大占据过多资源。从分析逻辑来看，关联分析无法进行统计检验，只能提示潜在相关关系，不能作为因果关系判定依据^[20]。

在机器学习领域，对样本进行分类是最常见的问题之一，可采用分类算法建立基于大样本、多变量的分类模型（分类器）完成对样本类别的预测。常用的分类算法包括

Logistic 回归、K 近邻、支持向量机、决策树、随机森林、人工神经网络^[37,46]等。在药品不良反应研究中，可根据是否发生不良反应对样本分组，以潜在影响因素为变量建立分类模型，根据最终模型各特征重要性来评价与不良反应相关的风险因素。刘欢^[47]基于自发呈报系统数据库对参麦注射液不良反应发生因素进行研究，首先通过统计检验初步筛选出与不良反应发生和预后相关的个体因素、疾病因素、药物因素，分别以 Logistic 回归、决策树、支持向量机和人工神经网络为算法建立分类模型。选择出其中预测准确性最好的 Logistic 回归、支持向量机、人工神经网络这 3 个模型继续优化，最终根据模型变量重要性指标评估出对参麦注射液不良反应的影响因素。

在药品不良反应监测领域，除了对不良反应信号进行挖掘，以及评估引发不良反应的风险因素研究外，建立不良反应预测机制，即根据识别出来的潜在因素，预测药品发生不良反应的概率，也是未来的研究方向。张佳颖^[48]针对血栓通注射液和血塞通注射液临床不良反应监测，首先识别出患者因素、用药因素等 53 个变量，采用特征选择方法从中筛选 10 个最重要的特征，分别采用 K 近邻、支持向量机、梯度提升决策树、随机森林建立不良反应分类模型。最终，将其中表现最优的梯度提升决策树分类器部署进 web 应用框架中，构建了不良反应前置预测系统，在临床不良反应管控中具有实用价值。

3 数据挖掘应用于中成药安全性评价问题探讨

随着信息化医疗服务系统的完善，以电子医疗数据为资源，借助数据挖掘技术对中成药不良反应进行研究，已经成为中成药上市后安全性评价发展趋势。但是，该技术在中成药不良反应监测领域的推广仍然面临诸多挑战。

3.1 建立医疗大数据共享机制 开展数据发掘研究需要大量来源丰富的医疗数据。一般单个医院的医疗数据库存量相对较小，需要将不同医院，甚至不同地区的医疗数据进行汇总分析，才能更加全面地对中成药安全性进行评价。但是，来自不同省市医院电子医疗系统数据的记录内容难以达到完全一致，或者存在数据缺失等现象，同时出于对患者信息保密等原因，无法将各个医院数据进行汇总与融合，使数据挖掘技术对大规模数据分析的优势难以充分发挥，得到的结果不能全面覆盖潜在的风险因素。针对该问题，有研究建议借鉴 AsPEN 分布式网络的分析模式，建立协调中心指导各地采用自己数据库独立完成数据分析，由协调中心对各个结果汇总并得到全国的结果，使得研究结果更具有外推性^[36]。也有研究倡导首先以顶层设计方式构建统一的数据结构与变量，前瞻性地多个医院收集数据，最终形成严格设计与标准化的前瞻性医院信息数据库，对该数据库开展数据挖掘研究^[4]。

3.2 辨识中成药不良反应物质基础 中成药化学成分复杂，上市前药理以及毒理研究难以对各种化学成分性质进行充分解析，导致引起潜在不良反应的物质基础不明确。药品质量、患者状况以及合并用药等因素均有可能与不良

反应产生相关性，导致引发中成药不良反应的成因非常复杂^[49]。其中，引起中成药不良反应的潜在质量因素是制药企业及药品监管机构重点关注的问题。中成药与质量相关的化学成分可以看作庞大的变量集群，数据挖掘能够针对庞大而复杂变量进行深度解析，发现变量之间的相关关系和因果关系，已经有报道采用数据挖掘技术开展中成药生产过程及质量评价研究^[50-51]。但是，借助数据挖掘来辨析导致中成药不良反应的化学物质基础的研究罕有报道。随着现代分析技术的进步，中成药质量数据采集技术越发多样化，如色谱指纹图谱^[52]、质谱^[53]、定量核磁^[54]等。可以将这些质量数据与临床用药数据，以及不良反应信息进行整合，形成规模更大、信息维度更加丰富的数据系统。借助数据挖掘技术对该系统进行深入分析，发现与不良反应相关联的化学物质信号，为从复杂中成药化学成分中寻找引起不良反应的物质提供线索。

3.3 培养中成药数据挖掘人才队伍 数据挖掘是综合了计算机、数学、统计学、信息学等多个学科的综合技术，涉及的理论十分复杂，专业性很强。传统中医药领域缺乏数据挖掘专业人才，随着中成药不良反应监管要求的严格，这方面人才缺口将进一步扩大。因此，有必要加强中药领域数据挖掘技术人才的培养，鼓励学习与掌握先进的大数据技术理论与方法，促进传统中成药临床安全性评价与创新方法的融合。近年来，大数据、机器学习等技术逐渐得到中医药行业的重视，开始向中药不同领域的研究中渗透^[55-56]。但是，针对中成药不良反应开展的数据挖掘研究多是由高校或研究机构主导，医院提供丰富的数据资源，中成药生产企业在这方面的研究较少。药品安全性是中成药生产企业肩负的使命，企业应当发挥主体责任，积极开展这方面的研究工作。有必要加强医疗机构、研究院所与药品生产企业之间的合作，通过将各自所掌握的数据共享，让中成药生产企业参与到上市后不良反应监测的研究项目中来。一方面提升药品生产企业对数据挖掘技术人才队伍的建设；另一方面促进药品生产企业增强对自身产品的认识，从而提高对药品质量的管控能力。

4 结语与展望

信息化医疗系统的发展为中成药临床安全性研究提供了大数据资源，电子医疗数据库具有数据量大、信息丰富、来源真实等优势。随着大数据技术的兴起，借助数据挖掘创新理念及其丰富的算法，可用于发现中成药不良反应信号，筛选引发中成药不良反应的相关因素，在中成药安全性评价研究中显示出良好的应用前景。但是，该方法的大规模推广仍面临诸多挑战，例如需要加强多中心医疗大数据的共享，加快中医药领域数据挖掘人才队伍的建设，鼓励中成药生产企业参与不良反应监测数据分析等。另一方面，中成药是一种复杂物质体系，引发中成药不良反应的化学物质基础不明确，如何借助数据挖掘技术辨析引起不良反应的化学成分来源，是未来中成药安全性评价领域的研究方向之一。

中成药安全性是中医药事业健康发展的核心。开展药品不良反应监测研究,明确中成药不良反应事项,不仅是医疗机构和药品监管部门关心的问题,也成为中药生产企业产品存续的前提条件。数据挖掘作为信息化和大数据时代最有竞争力的技术之一,在中成药不良反应监测研究中显示出了良好的应用前景。随着数据挖掘算法的不断进步,以及中成药化学成分及毒理机制研究的深入,促进数据挖掘技术与中药自身特点、临床不良反应信息深度融合应用,对于探索中成药不良反应发生规律,预防药品不良事件,指导临床合理用药具有现实意义。

参考文献:

[1] 国家药品不良反应监测中心. 国家药品不良反应监测年度报告 (2020 年) [EB/OL]. (2021-03-26) [2023-06-05]. https://www.cdr-adr.org.cn/tzgg_home/202103/t20210326_48414.html.

[2] 国家药品不良反应监测中心. 国家药品不良反应监测年度报告 (2021 年) [EB/OL]. (2022-03-30) [2023-06-05]. https://www.cdr-adr.org.cn/drug_1/aqjs_1/drug_aqjs_sjbg/202203/t20220330_49586.html.

[3] 国家药品不良反应监测中心. 国家药品不良反应监测年度报告 (2022 年) [EB/OL]. (2023-03-24) [2023-06-05]. https://www.cdr-adr.org.cn/drug_1/aqjs_1/drug_aqjs_sjbg/202303/t20230324_50019.html.

[4] 吕建,王连心,谢雁鸣. 基于处方序列与处方序列对称分析的中药药物警戒[J]. 中国中药杂志, 2021, 46(21): 5468-5474.

[5] 国家药品监督管理局. 国家药监局关于发布《中药注册管理专门规定》的公告 (2023 年第 20 号) [EB/OL]. (2023-02-10) [2023-06-05]. <https://www.nmpa.gov.cn/xxgk/fgwj/xzhgfxwj/20230210173401120.html>.

[6] Ribeiro-Vaz I, Santos C C, Cruz-Correia R. Promoting adverse drug reaction reporting: comparison of different approaches[J]. *Rev Saude Publica*, 2016, 50: 14.

[7] Stang P E, Ryan P B, Racoosin J A, *et al*. Advancing the science for active surveillance: rationale and design for the Observational Medical Outcomes Partnership [J]. *Ann Intern Med*, 2010, 153(9): 600-606.

[8] 李海龙,赵厚宇,周一帆,等. 基于电子医疗数据库的药品不良反应信号挖掘方法概述[J]. 药物流行病学杂志, 2018, 27(8): 541-549.

[9] 王玲,周晶鑫,张学宁,等. 药品不良反应报告数据规整的实践与思考[J]. 中国药物警戒, 2014, 11(7): 407-412.

[10] 廖星,于丹丹,谢雁鸣,等. 碟脉灵苦碟子注射液上市后安全性监测:一项 30233 例真实世界注册登记研究[J]. 中国中药杂志, 2017, 42(15): 2857-2863.

[11] 姜俊杰,谢雁鸣,张寅,等. 30209 例舒血宁注射液安全性医院集中监测研究[J]. 中国中药杂志, 2017, 42(15): 2883-2888.

[12] 王志飞,喻锦扬,谢雁鸣. 参附注射液 30106 例上市后临床

安全性医院集中监测[J]. 中国中药杂志, 2017, 42(15): 2871-2876.

[13] 黎元元,方邦江,刘健,等. 30013 例悦安欣苦碟子注射液安全性医院集中监测[J]. 世界中医药, 2020, 15(1): 30-34.

[14] 谢雁鸣,廖星,姜俊杰,等. 中成药上市后安全性医院集中监测技术规范[J]. 中国中药杂志, 2019, 44(14): 2896-2901.

[15] 冯硕,刘文佳,胡晶,等. 药品安全性评价技术及其在中药领域的应用[J]. 中国合理用药探索, 2020, 17(1): 28-31.

[16] 王玲. 试论医疗大数据给药品安全性监测与评价带来的机遇和挑战[J]. 中国药物警戒, 2015, 12(11): 660-664.

[17] Andrews E B, Margulis A V, Tennis P, *et al*. Opportunities and challenges in using epidemiologic methods to monitor drug safety in the era of large automated health database[J]. *Curr Epidemiol Rep*, 2014, 1(4): 194-205.

[18] 杨羽,詹思延. 上市后大数据药品安全主动监测模式研究的必要性和可行性[J]. 药物流行病学杂志, 2016, 25(7): 401-404; 413.

[19] 程翼宇,钱忠直,张伯礼. 创建以过程管控为核心的中药质量控制技术体系[J]. 中国中药杂志, 2017, 42(1): 1-5.

[20] 张寅,谢雁鸣,陈岑,等. 基于关联规则 Apriori 算法的真实世界复方苦参注射液治疗恶性肿瘤联合用药药理作用特征的回顾分析[J]. 中国中药杂志, 2017, 42(2): 378-384.

[21] Han J, Kamber M. 数据挖掘: 概念与技术[M]. 北京: 机械工业出版社, 2010.

[22] 高嵩,高月娟,朱仁英,等. 数据挖掘技术在药品不良反应监测中的应用进展[J]. 中国当代医药, 2021, 28(26): 31-35.

[23] 刘梦玲,章新友,丁亮,等. 数据挖掘方法在中药配伍规律研究中的应用与进展[J]. 中国中药杂志, 2021, 46(20): 5233-5239.

[24] 张胜,陈厚柳,瞿海斌. 中药产品质量回顾数据挖掘方法研究[J]. 中国中药杂志, 2023, 48(5): 1264-1272.

[25] 李娟娟. 药品不良反应信号检测方法理论及应用研究[D]. 西安: 第四军医大学, 2008.

[26] 夏旭东,程伟高,周明,等. 基于真实世界数据的清开灵注射液不良反应信号挖掘与分析[J]. 中草药, 2023, 54(4): 1220-1227.

[27] 陈炯华,魏永越,孙骏,等. 基于数据挖掘的预警方法在双黄连注射剂不良反应监测中的应用研究[J]. 中国中药杂志, 2010, 35(3): 308-312.

[28] 向永洋,谢雁鸣,易丹辉. 药物警戒数据挖掘方法比较及其在中成药预警中的应用[J]. 中国中药杂志, 2011, 36(20): 2831-2835.

[29] Petri H, De Vet HC, Naus J, *et al*. Prescription sequence analysis: a new and fast method for assessing certain adverse reactions of prescription drugs in large population[J]. *Stat Med*, 1988, 7(11): 1171-1175.

[30] 艾青华,曾宪斌,谢雁鸣,等. PSA 结合 NCCS 研究参芪扶

正注射液疑似过敏反应影响因素[J]. 中国中药杂志, 2014, 39(18): 3563-3566.

[31] 白 洋, 杨 薇, 刘 岷, 等. 734 例参附注射液疑似过敏反应影响因素分析[J]. 中成药, 2016, 38(3): 505-510.

[32] 李 敏, 谢雁鸣, 章秩立, 等. 基于巢氏病例对照研究丹参川芎嗪注射液疑似过敏反应的影响因素分析[J]. 中国中医基础医学杂志, 2019, 25(12): 1707-1709.

[33] Hallas J. Evidence of depression provoked by cardiovascular medication: a prescription sequence symmetry analysis [J]. *Epidemiol*, 1996, 7(5): 478-484.

[34] 陶庆梅, 詹思延. 处方序列分析与处方序列对称分析在药物流行病学中的应用[J]. 药物流行病学杂志, 2012, 21(10): 517-519.

[35] 周一帆, 程吟楚, 高双庆, 等. 以喜炎平注射液为例探究处方序列对称分析对过敏信号的检出能力[J]. 药物流行病学杂志, 2019, 28(6): 376-379; 388.

[36] 胡 晶, 冯 硕, 李 博, 等. 处方序列对称分析在研究中药注射剂不良反应中的应用展望[J]. 世界科学技术 (中医药现代化), 2018, 20(10): 1739-1744.

[37] Geron A. 机器学习实战 [M]. 北京: 机械工业出版社, 2020.

[38] 吉萌萌. 清开灵注射液医院集中监测不良反应影响因素研究[D]. 北京: 中国中医科学院, 2020.

[39] 王惠霞. 基于 Logistic 回归分析的活血化瘀类中药注射剂药品不良反应/药品不良事件影响因素分析[J]. 中国中医药信息杂志, 2019, 26(7): 133-136.

[40] 张广辉, 孙粼希, 吴显文, 等. 巢氏病例对照联合 SMOTE 算法分析注射用灯盏花素肾功能异常影响因素[J]. 中国药物警戒, 2021, 18(12): 1150-1154.

[41] 周 虎. 基于电子病历的痰热清注射液疑似过敏反应信号发现与影响因素探讨[D]. 重庆: 重庆医科大学, 2020.

[42] 郭 鑫, 裴贵珍, 鲁宁杰, 等. 基于真实世界丹参川芎嗪注射液不良反应信号挖掘[J]. 安徽医药, 2020, 24(2): 421-424.

[43] 张文辉, 赵文光. 基于数据挖掘的药物不良反应因果关系研究[J]. 中国数字医药学, 2019, 14(5): 43-45.

[44] 高 强, 秦 柯, 王月峰. 热毒宁注射液不良反应影响因

素的关联规则分析[J]. 现代药物与临床, 2023, 38(4): 962-968.

[45] 崔盈盈, 吴嘉瑞, 谭 迪, 等. 基于关联规则的双黄连注射剂不良反应流行病学特点研究[J]. 药物流行病学杂志, 2018, 27(2): 85-91; 130.

[46] 程翼宇, 瞿海斌. 医药分析信息学及分析数据处理技术[M]. 北京: 化学工业出版社, 2006.

[47] 刘 欢. 基于巢式病例对照的参麦注射液不良反应的数据挖掘分析[D]. 成都: 电子科技大学, 2018.

[48] 张佳颖. 以含三七皂苷成分中药注射液为例建立 ADR 前置预测系统[D]. 成都: 电子科技大学, 2020.

[49] 童元元, 杨金生, 李长虹, 等. 基于数据挖掘的中药安全性研究与思考[J]. 中国中药杂志, 2011, 36(21): 3047-3050.

[50] Zhang S, Xie X Y, Qu H B. A data-driven workflow for evaporation performance degradation analysis: a full-scale case study in the herbal medicine manufacturing industry[J]. *J Intell Manuf*, 2023, 34: 651-668.

[51] 孙莹莹, 孙 巍, 王 萍, 等. 基于 UPLC 指纹图谱的逍遥片批次一致性评价研究[J]. 中草药, 2020, 51(20): 5154-5165.

[52] 田清华, 张 亚, 徐 聪, 等. HPLC 指纹图谱结合化学模式识别评价舒筋活血制剂的整体质量[J]. 中草药, 2019, 50(20): 4925-4931.

[53] Tang J F, Li W X, Tan X J, *et al.* A novel and improved UHPLC-QTOF/MS method for the rapid analysis of the chemical constituents of Danhong injection [J]. *Anal Met*, 2016, 8: 2904-2914.

[54] 李文竹, 谢欣媛, 杨嘉誉, 等. 同时测定丹红注射液中 21 个化学成分含量的 NMR 方法研究[J]. 中国中药杂志, 2022, 47(23): 6399-6408.

[55] 李 正, 康立源, 范骁辉. 中药制药过程数据集成、数据挖掘与可视化技术研究[J]. 中国中药杂志, 2014, 39(15): 2989-2992.

[56] 程翼宇, 张伯礼, 方同华, 等. 智慧精益制药工程理论及其中药工业化转化研究[J]. 中国中药杂志, 2019, 44(23): 5017-5021.