

研究报告

毒理学数据可靠性评价工具在谷氨酸及其盐神经毒性危害评估中应用

梁婕¹,俞萍²,谢文静³,吕中明²,周永林²,卞倩^{1,2}

(1. 南京医科大学公共卫生学院, 江苏 南京 211166; 2. 江苏省疾病预防控制中心, 江苏 南京 210009; 3. 东南大学公共卫生学院, 江苏 南京 210009)

摘要:目的 评价谷氨酸及其盐神经毒性数据的可靠性并给出使用建议,为谷氨酸及其盐类危害评估提供科学保证;在使用过程中检验毒理学数据可靠性评价工具(toxicology data reliability assessment method, TRAM),为进一步完善工具提供参考。方法 使用TRAM对通过系统综述中检索文献的方法筛选出的60篇谷氨酸及其盐神经毒性相关文献进行可靠性评价。根据文献所涉及的数据类型(动物试验数据、人类数据)进入相应的评价流程,按照评分要素及建议分值对各个评价指标逐条打分。计算每条得分与对应权重的乘积之和,除以各类数据总分得出质量分数,最后根据质量分数判定可靠性等级并给出使用建议。其中,每篇文献的评价工作均由相关领域的两名人员独立完成。结果 最终12篇文献可靠性等级评价为“高”,建议优先使用;43篇文献可靠性等级评价为“中”,可以使用;5篇文献可靠性等级评价为“低”,不建议使用。结论 TRAM的评价指标兼顾了报告质量和方法学质量,且填补了其他工具涉及人类毒理学资料可靠性评价方法的空白。TRAM对食品安全风险评估工作具有更好的适用性,为更加客观、科学地开展危害识别以及风险评估工作提供了方法和手段。

关键词:毒理学数据;可靠性;毒理学数据可靠性评价工具;谷氨酸;危害评估

中图分类号:R155 文献标识码:A 文章编号:1004-8456(2020)05-0509-05

DOI:10.13590/j.cjfh.2020.05.007

Application of toxicology data reliability assessment method, a toxicological data reliability evaluation tool, in the neurotoxic hazard assessment of glutamate acid and its salts

LIANG Jie¹, YU Ping², XIE Wenjing³, LYU Zhongming²,
ZHOU Yonglin², BIAN Qian^{1,2}

(1. School of Public Health, Nanjing Medical University, Jiangsu Nanjing 211166, China;

2. Center for Disease Control and Prevention of Jiangsu Province, Jiangsu Nanjing 210009, China;

3. School of Public Health, Southeast University, Jiangsu Nanjing 210009, China)

Abstract: Objective This paper aims to evaluate data reliability of the neurotoxic hazard assessment of glutamate and its salts and provide recommendations, as well as to improve toxicology data reliability assessment method (TRAM) via trial application. **Methods** TRAM was used to evaluate the reliability of 60 articles which were selected by the method of systematic review documentation retrieval. The evaluation was based on types of toxicological data involved in each paper (laboratory animal data or human data) and they were scored by reliability criteria. The quality percentage was obtained via calculations to judge reliability categories and give recommendations. It's necessary to note that the evaluation of each paper was independently completed by two persons in related fields. **Results** After three rounds of evaluation, the reliability of 12 articles were evaluated as “high” and recommended for priority use. The reliability of 43 articles was rated as “moderate” and can be used. The reliability of 5 articles was evaluated as “low” and not recommended to use. **Conclusion** TRAM takes both reporting quality and methodological quality into consideration, especially including human data reliability evaluation method which is absent in the other toxicology data reliability assessment tools. TRAM is more suitable for food safety risk assessment. It provides a better objective and scientific guarantee for hazard identification and risk assessment.

Key words: Toxicological data; reliability; toxicology data reliability assessment method; glutamate; hazard assessment

谷氨酸及其盐由于在食品加工过程中具有增

鲜作用而被我国与其他国家和地区批准作为食品添加剂,并被广泛使用。我国是世界上谷氨酸钠(味精的主要成分)最大的生产、消费和出口国之一^[1],有研究报道谷氨酸钠的摄入可能与肥胖、代谢综合征、中餐馆综合征等健康风险相关^[2]。至今,关于谷氨酸及其盐神经毒性方面的研究颇多,

收稿日期:2020-06-01

作者简介:梁婕 女 硕士生 研究方向为卫生毒理学 E-mail: starryliangjie@163.com

通信作者:卞倩 女 主任医师 研究方向为毒理学与风险评估 E-mail: bianqian@jcdc.cn

但结论并不一致。2017年,欧洲食品安全局(European Food Safety Authority, EFSA)对谷氨酸及其盐的健康风险进行了再评估,根据1979年的一项大鼠神经发育毒性研究结果建立谷氨酸及其盐的每日容许摄入量(acceptable daily intake, ADI)为30 mg/kg BW(不确定系数为100倍)^[3],但该研究者于2018年再次针对该试验结果发表了论述,文章系统地回顾和分析了1979年的试验设计、试验方法及数据的分析,指出1979年的研究并没有为谷氨酸钠的神经发育毒性提供有说服力的证据,认为该研究结果并不能为未观察到有害作用水平(no observed adverse effect level, NOAEL)的确定及修改谷氨酸钠的ADI值提供科学基础^[4]。基于上述原因,全面评估现有谷氨酸及其盐健康风险相关研究证据,为我国谷氨酸及其盐的管理提供科学依据至关重要。

1 资料与方法

1.1 文献筛选

在对数据进行可靠性评价之前,课题组首先用系统综述检索文献的流程进行文献^[5]筛选。以“monosodium glutamate (monosodium L-glutamate, 谷氨酸钠), neurotoxicity (brain damage, developmental neurotoxicity, 神经毒性, 脑损伤, 神经发育毒性), oral (digestion, 经口)”为主要检索关键词从Pubmed、Web of Science、Embase、Toxline等数据库中检索2018年10月14日以前发表的文献,并人工检索了相关主题综述或报告,共检索到9 030篇文献,通过阅读摘要排除了2 892篇重复文献。根据纳入和排除标准对剩余的6 137篇文献进行筛选。其中,文献纳入标准为:(1)原创性研究;(2)中英文文献;(3)文献内容是关于谷氨酸及其盐对动物个体的影响;(4)文献内容是关于谷氨酸及其盐通过消化道途径对动物的影响;(5)文献的内容是关于谷氨酸及其盐对动物的神经毒性影响;(6)文献的内容是关于谷氨酸及其盐对机体健康效应的影响。文献排除标准为:(1)病例报告、评论、社论、综述、信件或者报告方面的文献;(2)除中英文外的其他语言文献;(3)文献内容是谷氨酸及其盐对细胞等的影响;(4)文献内容是谷氨酸及其盐通过其他非消化道途径,如腹腔内注射、皮下注射等对动物的影响;(5)文献内容是谷氨酸及其盐导致细胞毒性等其他毒性影响;(6)文献内容是不包括其他物质对机体健康影响的文献。最终纳入了60篇文献进行可靠性评价。其中2篇文献同时包含人类数据和动物试验数据,其他58篇均为动物试验数据。

1.2 评价流程

毒理学数据可靠性评价体系TRAM的评价流程如图1所示。

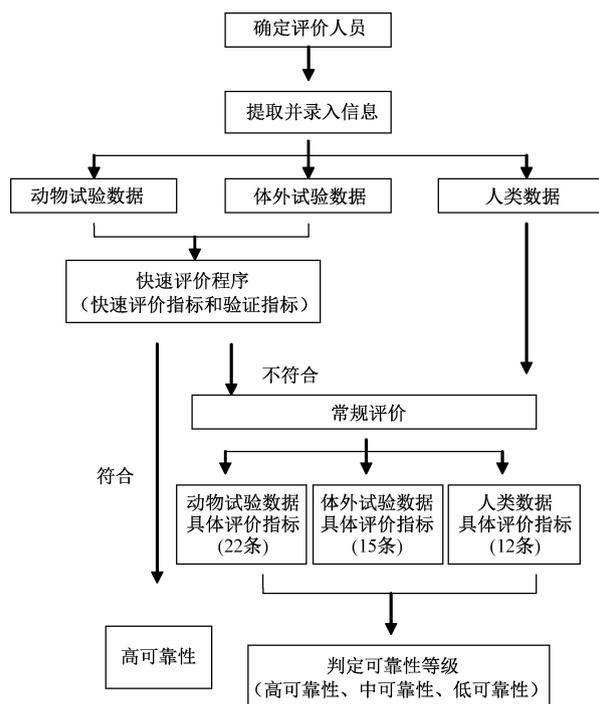


图1 毒理学数据可靠性评价体系TRAM的评价流程图

Figure 1 Toxicology data reliability assessment method flow chart

1.2.1 评价人员组成

由课题组11名评价人员组成可靠性评价小组,所有评价者均有5年以上毒理学工作经验,其中高级职称8人。高级职称中有3人参与了TRAM的方法学构建,6人曾参与了该评价体系的方法比对与验证工作,包括使用同一工具评价不同文献以及使用不同方法评价同一文献等。所有人员在接受TRAM培训并经过模拟评价,能熟练使用该工具后对最终纳入的60篇文献进行评价。

1.2.2 确定评价人员

根据每篇待评价文献的数据类型(动物试验数据、人类数据),确定2名具有合适专业背景的人员进行评价工作。

1.2.3 录入评价人员和数据基本信息

输入评价人员的姓名、职称、工作单位等基本信息,以及毒理学数据的作者姓名、题目、发表年代、作者所在机构等基本信息。

1.2.4 根据数据类型选择相应的评价通道

根据待评价数据的类型选择进入快速评价通道或常规评价通道进行评价。

1.2.5 评价的基本过程

评价过程共经历3轮。第一轮,每篇待评价文献分别由2名评价者使用TRAM独立完成评价。

第二轮,对于两个评价结果可靠性等级不同或评分差距在 10 分以上文献,由两位原评价人员就主要分歧充分讨论后分别重新打分。若经第二轮评价后双方仍不能得到统一的可靠性等级则进入第三轮评价,由第三位专业人员或相关专家介入,并以少数服从多数的原则确定最终的评价等级。

1.3 评价内容

1.3.1 快速评价

对于动物试验数据(60 篇待评价文献均不涉及体外试验数据,体外试验数据的证据等级相对偏低,只有在动物试验数据和人类数据不充分的情况下才会使用体外试验数据),首先进入快速评价通道进行评价。符合快速评价指标且满足验证指标的直接评价为高可靠性等级,否则进入具体指标评价流程。

快速评价指标必须同时符合以下两个要求:(1)试验按照国际公认或国内统一的试验程序开展[经济合作与发展组织(OECD)、欧盟(EU)、美国食品药品监督管理局(FDA)等的试验指南,我国国家标准、技术规范等];(2)数据来源于遵循良好实验室规范(GLP)、中国合格评定国家认可委员会(CNAS)或 ISO/IEC 17025 等质量管理体系并获得相应的国内、国际权威机构认可的实验室(详见表 1)。

表 1 快速评价验证指标

Table 1 Validation indicators for rapid evaluation

| 受试物必须能被识别 | 试验结果完整 |
|--|--|
| 1. 动植物、微生物受试物提供拉丁名 | |
| 2. 单一组受试物提供 CAS 号(其他类似 CAS 的唯一性编号)和化学结构式 | 提供了完整的试验结果 |
| 3. 混合物提供了其组成成分,且符合条款 2 | |
| 4. 评价人员认为受试物能被准确识别,即受试物具有唯一性,不会与其他物质相混淆 | 例如:是否提供了各观察终点的结果,包括对照在内的结果和细节,是否有数据缺失,阴性结果也应报告 |

1.3.2 常规评价

人类数据及不符合快速评价程序的动物试验数据进入相应的具体评价,按照评价指标所列出的评分要素及建议分值逐条打分。动物试验数据分为受试物、实验动物、试验设计、试验操作与方法、试验结果和结论 5 部分共 22 条具体评价指标,人类数据分为研究类型和研究对象、测量、质量控制、偏倚控制、结果和结论 5 部分共 12 条具体评价指标,所有指标满分均为 4 分,对应不同权重。动物试验数据及人类数据具体评价指标及权重参照文献[6]。

1.3.3 计算实际得分与质量分数

计算每条具体评价指标得分与所对应权重的乘积之和,得到实际得分(实际得分 = \sum 每条评价指标得分 \times 各评价指标的权重)。用实际得分除以各类型数据总分(动物试验数据总分 = $\sum_{k=1}^{22} 4$ 分 \times 各评价指标的权重;人类数据总分 = $\sum_{k=1}^{12} 4$ 分 \times 各评价指标的权重)得出百分比,即质量分数。

1.3.4 判定可靠性等级及使用建议

根据质量分数判定数据的可靠性等级,并给出相应的使用建议,具体见表 2。

表 2 动物和人类数据可靠性等级判定及使用建议

Table 2 Animal and human data reliability categories

| 实际得分(总分) | 可靠性等级 | 使用建议 |
|----------|-------|------|
| <60% | 低 | 不使用 |
| 60%~80% | 中 | 可以使用 |
| >80% | 高 | 优先使用 |

注:可靠性等级为低的数据原则上不用于食品安全风险评估,可靠性等级为中和高的数据可以用于食品安全风险评估,建议优先使用可靠性等级为高的数据

2 结果

2.1 第一轮可靠性评价结果

第一轮评价,每篇文献均由两名评价人员分别独立评价,结果 20 篇文献评价结果等级不一致或评分差距在 10 分以上,评分差距最大的文章分值相差 33.7 分。40 篇文献评价等级一致或评分差距在 10 分以内,评价结果等级一致率为 66.7%。

2.2 第二轮可靠性评价后的结果

第二轮评价,对 20 篇评价结果等级不一致或评分差距在 10 分以上的文献进行讨论后分别重新打分,最终只有 1 篇文献评价结果等级不一致,评价分值相差为 10.5。经过第二轮评价,评价结果等级一致率为 98.3%。

2.3 最终评价结果

经过第三位相关专家的介入,最终确定了第二轮评价中仍未得到同一等级的那篇文献的评价等级。最终可靠性等级评价结果见表 3,其中 12 篇可靠性等级为“高”和 43 篇可靠性为“中”的文献被纳入。其中可靠性等级为“高”的 12 篇文献 22 项具体评价指标得分分值见表 4,每项指标满分均为 4 分,依据相应的评分要素及建议分值给出 0~4 分。

表 3 文献数据的可靠性等级评价结果

Table 3 Reliability evaluation results of literature data

| 可靠性等级评价结果 | 高 | 中 | 低 | 合计 |
|-----------|----|----|---|----|
| 动物试验数据 | 12 | 44 | 4 | 60 |
| 人类数据 | 0 | 1 | 1 | 2 |
| 文献可靠性 | 12 | 43 | 5 | 60 |

注:共 60 篇文献,其中 58 篇为动物试验数据,2 篇同时包含人类数据和动物试验数据

表4 可靠性等级评价为“高”的12篇文献每项指标评分结果
Table 4 Score results of each indicator of 12“high” reliability articles

| 评价指标 | 得分 | | | | | | | | | | | |
|---------------------------|------|------|------|------|------|------|------|------|------|------|------|------|
| | 文献1 | 文献2 | 文献3 | 文献4 | 文献5 | 文献6 | 文献7 | 文献8 | 文献9 | 文献10 | 文献11 | 文献12 |
| 1. 受试物的识别 | 4 | 4 | 4 | 3 | 4 | 2 | 4 | 4 | 4 | 2 | 4 | 2 |
| 2. 受试物的特征描述 | 0 | 2 | 2 | 2 | 0 | 4 | 2 | 0 | 4 | 2 | 2 | 2 |
| 3. 受试物的来源 | 2 | 2 | 2 | 4 | 0 | 0 | 2 | 0 | 0 | 2 | 4 | 2 |
| 4. 受试物配置 | 4 | 4 | 0 | 3 | 2 | 4 | 3 | 0 | 4 | 0 | 4 | 2 |
| 5. 实验动物的物种 | 4 | 4 | 4 | 4 | 4 | 4 | 4 | 4 | 4 | 4 | 4 | 4 |
| 6. 实验动物的来源 | 2 | 4 | 4 | 4 | 2 | 0 | 0 | 4 | 0 | 2 | 4 | 2 |
| 7. 实验动物的微生物级别 | 0 | 0 | 0 | 4 | 0 | 0 | 0 | 0 | 0 | 4 | 0 | 0 |
| 8. 实验动物的品系、遗传特点 | 0 | 4 | 4 | 4 | 4 | 4 | 4 | 4 | 4 | 4 | 4 | 4 |
| 9. 实验动物性别、初始体质量、年龄 | 2 | 4 | 4 | 3 | 4 | 4 | 4 | 4 | 4 | 4 | 4 | 4 |
| 10. 动物伦理福利 | 1 | 2 | 4 | 2 | 2 | 2 | 2 | 4 | 4 | 4 | 4 | 4 |
| 11. 实验动物饲养环境和饲养条件 | 1 | 2 | 0 | 4 | 2 | 0 | 2 | 4 | 2 | 2 | 2 | 4 |
| 12. 遵循的试验程序 | 4 | 0 | 4 | 3 | 4 | 4 | 2 | 0 | 4 | 4 | 4 | 4 |
| 13. 给样途径 | 4 | 4 | 4 | 4 | 4 | 4 | 4 | 4 | 4 | 4 | 4 | 4 |
| 14. 给样剂量的选择依据 | 4 | 4 | 4 | 3 | 4 | 4 | 4 | 4 | 4 | 4 | 4 | 3 |
| 15. 给样频率和持续时间 | 4 | 4 | 4 | 4 | 4 | 4 | 4 | 4 | 4 | 4 | 4 | 4 |
| 16. 对照 | 4 | 4 | 4 | 4 | 4 | 4 | 4 | 4 | 4 | 4 | 4 | 4 |
| 17. 实验动物数量和分组方法 | 4 | 4 | 2 | 3 | 4 | 4 | 4 | 4 | 2 | 4 | 4 | 4 |
| 18. 在经认可的质量管理规范或质量管理体系下试验 | 4 | 0 | 0 | 3 | 0 | 0 | 0 | 4 | 0 | 0 | 4 | 0 |
| 19. 统计方法 | 4 | 4 | 4 | 4 | 4 | 4 | 4 | 4 | 4 | 4 | 4 | 4 |
| 20. 缺失数据 | 4 | 4 | 4 | 4 | 4 | 4 | 4 | 4 | 4 | 0 | 4 | 4 |
| 21. 试验结果 | 4 | 4 | 4 | 4 | 4 | 4 | 4 | 4 | 4 | 4 | 4 | 4 |
| 22. 生物学效应 | 4 | 3 | 4 | 2 | 4 | 4 | 3 | 4 | 4 | 4 | 3 | 3 |
| 实际得分 | 176 | 172 | 172 | 187 | 174 | 174 | 174 | 168 | 180 | 170 | 192 | 170 |
| 质量分数/% | 84.6 | 82.7 | 82.7 | 85.0 | 83.7 | 83.7 | 83.7 | 80.8 | 86.5 | 81.7 | 92.3 | 81.7 |

注:12篇文献数据类型均为动物试验数据

3 讨论

毒理学数据是风险评估的核心和基础,而数据质量是风险评估结果的重要保证^[7]。以往的毒理学数据可靠性评价主要基于评价人员的经验值和主观判断,会导致同样的毒理学数据由不同的人员评价导致评价结果变异性增加,从而影响风险评估的科学性。ADAMSON^[8]曾开展了15份有关大鼠咖啡因急性半数致死量(LD₅₀)的研究(其中3项未发表),研究人员使用KLIMISCH等^[9]对已发表的12份研究数据的可靠性进行了评价,其中6份研究数据可靠性等级为不可靠类数据和无法归属类,通过对其余6份可靠性等级为可靠无限制类、可靠有限制类的研究数据进行分析,得到雄性大鼠LD₅₀的算数平均值为(367±37) mg/kg。而此前联合国发布的由未经质量评价的数据而得到的咖啡因急性LD₅₀值低于300 mg/kg。可见,用于风险评估的数据很有必要进行科学的可靠性评价。TRAM是针对食品安全风险评估中的毒理学数据而开发的可靠性评价工具,本次谷氨酸及其盐神经毒性再评估项目中对毒理学数据进行可靠性评价是TRAM在食品安全风险评估中的一次探索性应用。TRAM的评

价指标考虑了食物的膨胀系数,且重点关注经口、灌胃的给予途径,更适用于评估食品数据。风险评估工作涉及各种数据类型,而TRAM不仅包含动物试验数据和体外试验数据,还涵盖了以往的数据可靠性评价体系都未涉及的人类数据。人类数据较为珍贵且更为复杂,尤其是一些食物中毒报告、经过科学规范设计的人体试食试验和珍贵的食用史等,对于风险评估尤其是食品安全风险评估工作具有特殊的重要性,人类数据的加入对食品安全风险评估工作具有针对性更强、操作性更好的实际应用价值^[6-10]。KLIMISCH等^[9]评级系统由于指标不够明确和细化,评价结果变异性较大,ToxRTool中的评价指标主要关注毒理学数据的报告质量^[11],对试验设计合理性的评价关注不够^[12-13]。而TRAM弥补了上述两种评价工具的不足,它的评价指标兼顾了报告质量和方法学质量,评分标准也更加详细、客观,评价结果能更好地反映数据的可靠性。此外,TRAM中快速评价通道的设置极大地提高了评价的效率。TRAM曾与ToxRTool进行过对比验证工作,TRAM评价结果的高质量数据与ToxRTool评价结果的高质量数据基本一致(数据待发表)。

在对谷氨酸及其盐神经毒性危害评估项目数

据可靠性进行评价的过程中,项目组也发现 TRAM 指标中有不够完善的地方。比如,某些指标评分要点的描述不够清晰和明确,容易引起理解上的分歧;某些指标的主观性很大,容易对评分结果产生较大的影响;某些指标的建议分值划分笼统,可操作性差,容易造成评价人员无法根据评分标准给出客观分数的情况。经过课题组专家们反复的讨论,针对上述问题,对 TRAM 动物试验数据评价指标的建议分值部分进行了修改完善,如指标“9. 实验动物的初始体质量或年龄、性别”。原评分标准为:“提供了实验动物的性别、初始体质量、年龄得 4 分,未提供则不得分”。标准粗糙,未提及只提供了性别、年龄、体质量以及提供了三者中两者的情况该如何评分,可操作性差,根据提供信息的完整程度对评分指标进行细化后将指标 9 的建议分值调整为:“满分 4 分(需将条款 1 和 2 得分相加),条款 1. 提供了实验动物的初始体质量或年龄得 2 分,未提供则不得分。条款 2. 提供了实验动物的性别得 2 分,未提供则不得分”。经修改后,动物试验数据评价指标中酌情打分类别指标占比由 72.7% (16/22) 下降为 59.1% (13/22),客观评分类别指标占比由 27.3% (6/22) 上升至 40.9% (9/22),减少了评价的主观差异,增加了评价的客观科学性,且进一步明确和细化了评分标准,使得评分标准更加清晰、客观,便于评价人员操作。在 TRAM 后续的实践应用过程中,将逐步完善体外试验数据评价指标以及人类数据的评价指标。

TRAM 在谷氨酸及其盐危害评估中的应用是一个良好的开端。除了已经掌握毒理学信息的食品材料外,还有层出不穷的新的食品原料以及特殊食品材料,都需要进行食品安全风险评估,TRAM 将为更多的食品安全风险评估工作提供科学的数据可靠性保证。除了目前在食品安全风险评估项目中的应用,通过进一步的调整和完善,TRAM 还可以尝试应用于化妆品、一次性卫生用品等健康相关产品的风险评估工作中。此外,TRAM 还可以在科研工作中发挥作用,如通过对拟采用的参考文献进行评价,筛选出高可靠性的文献用于综述或系统综述,为科研选题、科研假说等提供更加科学、可靠的依据。同时,将 TRAM 应用于循证毒理学、系统毒理学以及计算毒理学等诸多领域,更好地为毒理学提供服务^[14-15]。

参考文献

- [1] KAZMI Z, FATIMA I, PERVEEN S, et al. Monosodium glutamate: review on clinical reports[J]. *Int J Food Prop*, 2017(20):1807-1815.
- [2] NIAZ K, ZAPLATIC E, SPOOR J. Guest editorial: extensive use of monosodium glutamate: a threat to public health? [J]. *EXCLI J*, 2018(17):964.
- [3] MORTENSEN A, AGUILAR F, CREBELLI R, et al. Re-evaluation of glutamic acid (E 620), sodium glutamate (E 621), potassium glutamate (E 622), calcium glutamate (E 623), ammonium glutamate (E 624) and magnesium glutamate (E 625) as food additives[J]. *EFSA J*, 2017,15(7):1-31.
- [4] VORHEES C V. A test of dietary monosodium glutamate developmental neurotoxicity in rats: a reappraisal[J]. *Annals of Nutrition and Metabolism*, 2018,73(5):36-42.
- [5] HOFFMANN S, DE VRIES R B M, STEPHENS M L, et al. A primer on systematic reviews in toxicology [J]. *Archives of Toxicology*, 2017,91(7):2551.
- [6] BIAN Q, YU P, WU J, et al. A new method to evaluate toxicological data reliability in risk assessments [J]. *Toxicology Letters*, 2019,311(5):125-132.
- [7] 李敏,徐海滨,何来英. 基于管理毒理学的毒性实验数据的系统评价[J]. *中国食品卫生杂志*, 2012,24(2):140-144.
- [8] ADAMSON R H. The acute lethal dose 50 (LD₅₀) of caffeine in albino rats[J]. *Regulatory Toxicology and Pharmacology*, 2016,80(7):274-276.
- [9] KLIMISCH H J, ANDREAE M, TILLMANN U. A systematic approach for evaluating the quality of experimental toxicological and ecotoxicological data [J]. *Regulatory Toxicology and Pharmacology*, 1997,25(1):1-5.
- [10] ROONEY A A, BOYLES A L, WOLFE M S, et al. Systematic review and evidence integration for literature-based environmental health science assessments [J]. *Environmental Health Perspectives*, 2014,122(7):711-718.
- [11] SCHNEIDER K, SCHWARZ M, BURKHOLDER I, et al. “ToxRTool”, a new tool to assess the reliability of toxicological data[J]. *Toxicology Letters*, 2009,189(2):138-144.
- [12] PRZYBYLAK K R, MADDEN J C, CRONIN M T D, et al. Assessing toxicological data quality: basic principles, existing schemes and current limitations [J]. *SAR and QSAR in Environmental Research*, 2012,23(5/6):435-459.
- [13] SEGAL D, MAKRIS S L, KRAFT A D, et al. Evaluation of the ToxRTool's ability to rate the reliability of toxicological data for human health hazard assessments[J]. *Regulatory Toxicology and Pharmacology*, 2015,72(1):94-101.
- [14] SILBERGELD E. Evidence-based toxicology: strait is the gate, but the road is worth taking[J]. *Altex*, 2013,30(1):67-73.
- [15] WATERS M D, FOSTEL J M. Toxicogenomics and systems toxicology: aims and prospects [J]. *Nature Reviews Genetics*, 2004,5(12):936.