

监督管理

基于管理毒理学的毒性实验数据的系统评价

李敏,徐海滨,何来英

(中国疾病预防控制中心营养与食品安全所,北京 100021)

摘要: 毒理学实验数据是反映化学品安全性的重要依据,在化学品的风险评估和风险管理中起着重要的作用。有效地利用不同实验室提供的毒理学数据,为管理毒理学服务,需要建立一套科学、可行的评价体系来评价毒理学数据的质量。目前国际上对毒理学数据评价系统的研究还刚刚起步,缺乏获得广泛共识的毒理学数据评价系统。本文简要论述了国际上现有的化学品毒理学数据评价系统的情况,为建立我国用于风险评估和 Risk 管理的毒理学数据评价系统提供思路。

关键词: 管理毒理学;实验数据;评价体系;质量;可靠性;相关性;充分性

中图分类号: R994.4 **文献标识码:** A **文章编号:** 1004-8456(2012)02-0140-05

Evaluation system of toxicology experiment data based on regulatory toxicology

Li Min, Xu Haibin, He Laiying

(National Institute for Nutrition and Food Safety, China CDC, Beijing 100021, China)

Abstract: Toxicological experiment data is the important evidence for the safety of chemicals and it plays an important role in the risk assessment and risk management of chemicals. In order to use the toxicological data from different laboratory to serve regulatory toxicology, it needs to establish a set of scientific and feasible criteria to evaluate the quality of experimental toxicological data. But the research about the systematic approach for evaluating the quality of toxicological data has just begun, and lack of a system reaching a broad consensus. This paper reviews the available international toxicological data evaluation system, which may help Chinese experts to establish an evaluation system in China.

Key words: Management toxicology; experimental data; evaluation system; quality; reliability; relevance; adequacy

化学品的安全性对于人体健康和环境安全是至关重要的。为了保证化学品特别是危险化学品的安全生产,防止化学品对人类健康和生态环境造成危害,从 20 世纪 80 年代开始,美国、欧盟和日本等发达国家已经普遍建立了各自的化学品管理法规体系。尽管各国的法规均是以化学品安全管理为核心,但在管理范围、管理内容以及管理强度上仍然存在着一定的差异。因此,为了促进国际化学品贸易的发展,避免各国因化学品管理要求的不同而发生贸易摩擦,联合国提出多项关于化学品安全管理的国际公约,以及新化学物质申报审批制度、化学品测试合格实验室和化学品风险评价等技术支持体系^[1]。进入 21 世纪以后,各国法规体系的建立和完善逐渐趋于国际化,化学品风险评价逐步成为化学品安全管理的重要方法,

国际组织和相关监管部门也一直致力于建立统一的全球通用的毒理学数据评估方法。

对化学品的危害和风险评估,主要基于化学品的各项实验数据,特别是毒理学和生态毒理学数据。然而,国际上对毒理学数据评价系统的研究还刚刚起步,缺乏获得广泛共识的毒理学数据评价系统。

近年来,我国法规监管部门也逐渐认识到化学品安全管理的重要性,参与了一些国际行动和国际公约,制定了符合我国国情的相关法律法规,并取得了一定的成效。但目前我国还没有建立系统的涉及化学品、药品、食品等普适性的毒理学资料质量的评价体系,而且欧盟等国际组织提出的化学品安全评估是一个相当广泛的概念,对于这一问题,我国还有很多实际问题需要解决。本文简要论述了国际上现存的有关化学品风险评估中毒理学数据评价系统的情况,为建立我国毒理学数据评价系统提供思路。

收稿日期:2011-08-23

基金项目:农业部行业工艺项目(200903054-06)

作者简介:李敏 女 硕士生 研究方向为食品安全 E-mail: xiaomi851019@163.com

通信作者:徐海滨 男 研究员 研究方向为食品安全
何来英 女 研究员 研究方向为食品安全

1 毒理学数据评价系统的概念

1.1 毒理学数据质量的差异

文献检索可以获得某种化学品毒性的大量数

据,包括单独的实验报告、刊物文献、综述、科普手册等。按照目前国际要求以及 REACH 法规对于数据的需求^[2],如果实验按照良好实验室规范 (GLP)或者其他国际承认的实验导则(如 EPA、OECD、EU、FDA 等实验操作指南)进行操作,那么所得的实验数据可以直接用于化学品风险评估中。然而检索到的诸多实验数据,难免有些数据的质量存在着较大的差异,参差不齐的数据对于化学品风险评估是否有参考价值,需要应用系统的评价程序来评价。

1.2 毒理学数据质量的内涵

评价数据的质量时应考虑 3 个方面:数据的可靠性、相关性和充分性。Klimisch 等^[3]对这 3 个方面做了如下描述:

1.2.1 数据的可靠性

评估实验报告和发表的文章是否按照标准的实验导则进行,对实验过程和结果的描述能够提出清晰而有说服力的证据。数据可靠性是评价数据质量的一个关键因素,它包括对实验方法、过程和结果的描述。上述几种毒理学数据评价系统主要是针对数据可靠性进行评价。

1.2.2 数据的相关性

用于描述实验数据适用于评估某种物质危害或风险程度的大小。在评价相关性时涉及以下相关问题^[4]:所测物质能否代表将要注册的物质,所选择的实验动物种群是否合适,试验中暴露途径与人群暴露途径是否相关等。理论上讲,人体实验数据对于某种化学品毒性评估来说是最具相关性的信息,但是人体实验数据毕竟有限,大多数还是动物实验、体外实验,所以常需要综合考虑这些数据,以便证明动物实验研究得到的结果是否与人相关^[5]。

1.2.3 数据的充分性

指数数据足够达到风险评估的目的,即数据在危害和风险评估中的可用性。评价实验数据的充分性需要考虑现有资料是否能够满足风险评估的需要,当现有数据不能满足评估需求时,需要专家确认“信息缺口”有多大,决定是否重新获取信息。

在评价数据的相关性和充分性时,总是会涉及到信息权重的问题^[6],例如当某一特定终点获得有多个实验的资料,这些资料有 1 个以上的结论时,需要应用权重程序来确定哪些结论可用于风险评估。最大的权重应该赋予最可靠和最相关的研究,这就需要为每个实验终点确定“关键研究”^[4]。通常这些“关键研究”是研究的质量、完整

性和代表性最好,与所评价的特异终点最具相关性的研究^[7]。

2 国际上主要的毒理学数据评价系统

国际上常用的毒理学数据评价系统有 3 种: Klimisch 评级系统^[3]、美国环境保护局信息筛选系统^[8]和毒理学数据可靠性评价工具 (ToxRTool)^[9]。

2.1 Klimisch 评级系统

由 Klimisch 等人于 1997 年提出,目前被欧洲化学品生态毒理学和毒理学中心 (European Centre for Ecotoxicology and Toxicology of Chemicals, ECETOC)、经济合作与发展组织 (Organization for Economic Cooperation and Development, OECD) 用于评价化学品毒理学实验数据的质量,该系统致力于对全球毒理学数据可靠性评价的一体化,同时也被广泛采纳于欧盟关于化学品注册、评估、授权和限制法令 (the Registration, Evaluation, Authorization and Restriction of Chemicals, REACH 法规) 卷宗整理中,可增强评价的透明度。

2.1.1 可靠性分类标准

Klimisch 评级系统将数据资料按可靠性分为 4 类^[3]:无限制的可靠类、有限制的可靠类、不可靠类和无法归属类。

2.1.2 分类方法

Klimisch 评级系统对数据的可靠性分类标准基本是按照 4 个类别的定义来制定的,例如,实验是否按照 GLP 或 OECD、EU 等规定的实验导则进行,严格按照导则实施的实验则被归为“无限制的可靠类”;没有按照上述导则实施的实验,则要进一步考虑其是否详细记录了实验过程和实验结果,如果记录完整,且有一定的科学性,则可归为“有限制的可靠类”;不符合的将其归为“不可靠类”或“无法归属类”。如此通过该评级系统将每一项数据归为不同的类别。通常情况下,第一、二类数据可用于危害风险评估,第三类“不可靠类”和第四类“无法归属类”的数据虽然一般不用于危害风险评估,但并非完全无用,在有些情况下可作为参考。

2.1.3 可靠性分类记录

对每一实验数据的分类要提供充分的依据,并且详细地记录下来,例如“该实验遵循 GLP 原则实施,所以归为无限制的可靠类”,或“该实验记录不充分,方法不具备科学性,所以归为不可靠类”等。通过这样详细的记录和分类使数据可靠性分类有证可循,体现了该评级系统的科学性,并且对于 REACH 应用 IUCLID 软件进行卷宗整理时更加方便^[10]。

Klimisch 评级系统将实验数据资料(特别是毒理学和生态毒理学资料的)可靠性分级标准化,自提出后被欧盟以及其他国家广泛应用,使后期的化学品风险评估工作更加方便快捷。

2.2 美国环境保护局信息筛选系统

美国环境保护局(Environmental Protection Agency, EPA)在1999年“高产量化学品挑战计划”中首次提出,用于评价可获得的化学品毒理学数据信息,并生成“筛选信息数据组(SIDS)”数据库^[11],作为化学品数据的检索工具为公众所用。

美国EPA为管理每年美国生产以及进口的高产量化学品(HPV),于1998年制定了“高产量化学品挑战计划”,目的是使公众能够方便地了解高产量化学品的危害和风险,以便更好地管理化学品。HPV是指每年在美国生产或进口量大于或等于100万磅的化学品。目前已经有2800种化学品列入该计划中。“HPV挑战计划”中也涉及到评价化学品数据信息质量的问题。EPA提出了“信息筛选系统”,通过这一系统筛选的信息将组成一个叫做“筛选信息数据组(SIDS)”的数据库,作为化学品毒性数据的检索工具为公众所用。

“信息筛选系统”由两层组成^[12]。关于某一特定化学品所能收集到的所有现存信息首先要进行第一层的初筛,通过预定的标准来评价数据信息的完整性,符合第一层标准的数据将进入第二层进行进一步筛选,不符合第一层要求的数据则不允许进入第二层。

对于该计划中的所有化学品的数据信息,最基本的要求是它的各项实验报告中都必须包括实验目的、实验方法的论述、结果和结论。第一层对于健康影响的毒理学数据的评价标准包括以下几个方面^[13]:实验动物物种、数量(每个实验剂量/浓度组的试验动物数,并按照性别分别记录);剂量/浓度水平(试验中受试物和对照物的剂量/浓度组数和剂量水平);暴露途径/方式(经口、皮肤、腹腔、吸入等);暴露时间;对照(所有试验都要有阴性对照组,有些试验要有阳性对照组);统计分析(必须描述所使用的统计分析方法,以及由该方法所得出的结果)。

第一阶段评价的目的就是依照上述标准对化学品现存信息进行初筛,数据信息要包含上述标准中的所有方面,如果有一项或多项不符合标准就不能进入第二阶段的评价,按照OECD的规定或相应的标准所得到的数据可直接进入第二阶段进行评价。进入第二阶段的数据信息将依照更为严格的标准进行进一步筛选。

Klimisch 评级系统和EPA信息筛选系统提出的时间较早,并且被很多国家和国际组织所使用。这两个系统虽然评价标准不尽相同,但是作用是一致的,都能够从大量的数据信息中筛选出可靠的信息,而去除不可靠的信息,不被EPA信息筛选系统所接受的数据同样不会被Klimisch所接受。例如,根据EPA信息筛选系统的要求,不符合上述所有标准的数据,不能进入第二阶段进行评价,而这些数据在Klimisch评级系统中属于不可靠类,也不能直接用于风险评估。另外,与Klimisch评价系统中的第三、四类数据一样,EPA信息筛选系统中未进入第二阶段的数据也可以作为参考资料用于风险评估中。所以,可以说这两个系统是并存的、相互兼容的,评价者在评价数据质量时可以单独使用或综合并用。

2.3 毒理学数据可靠性评价工具(ToxRTool)

欧洲替代方法验证中心(the European Centre for the Validation of Alternative Methods, ECVAM)于2008年开发了ToxRTool项目,旨在为评价现有毒理学数据建立一系列新的标准,使评价者的评价工作更加快捷。

2.3.1 ToxRTool的产生

虽然Klimisch评级系统历史悠久,应用广泛,但是数据分级标准仅仅是应用四个类别的定义来进行,评价者对于定义的理解可能有偏差,所以评价过程受评估者主观因素影响较大,缺乏更加客观的评价标准。因此ECVAM于2008年开发了一个研究项目,这个项目旨在为评价现有毒理学试验数据建立一系列新的更详细的标准,并且使评价者能够在方便、用户界面友好的工具中运用这些标准。

这个工具的主要目的是为评价者提供一个透明的方法将毒理学研究数据通过特殊的评价标准评价后分配到Klimisch 1、2、3类中。该工具只适用于原始数据,不适用于被归为Klimisch 4类的二手资料和手册等。

2.3.2 ToxRTool评价标准

该工具由两部分组成:体内实验标准和体外实验标准。这些标准最初是通过收集一系列对研究数据有潜在影响的参数而形成的,这些参数从评估指南和与研究相关的刊物以及报道中获得。经过反复研究与测试,ToxRTool的最终版本将标准设为5组,分别为:实验物质的鉴定、实验体系的特征、实验设计的描述、实验结果的记录、实验设计和结果的可靠性,并且细化体内实验为21条和体外实验为18条^[14],见表1。

表1 ToxRTool 数据可靠性评价标准
Table 1 ToxRTool evaluation standards for data reliability

分级标准	体内试验	体外试验
第1组:实验物质的鉴定	1. 受试物是否已被鉴别 2. 是否明确受试物的纯度 3. 受试物的来源是否明确 ^b 4. 对判断数据必不可少的每个测试项目的理化因素是否都已经明确 ^c	1. 受试物是否已被鉴别 2. 是否明确受试物的纯度 3. 受试物的来源是否明确 4. 对判断数据必不可少的每个测试项目的理化因素是否都已经明确
第2组:实验体系的特征	5. 实验动物的物种是否明确 ^d 6. 有机体的性别 7. 判定实验动物的血缘、种系 ^e 8. 实验开始时动物的年龄和体重 9. 只针对重复剂量毒理学研究:饲养条件 ^f	5. 是否描述了实验体系 6. 实验体系的来源信息是否明确(例如提供细胞株的实验室、实验体系供应商,体外实验器官组织细胞的来源等) 7. 有关实验体系的性质,培养条件等信息是否明确
第3组:实验设计的描述	10. 给药的途径 11. 给予的剂量或浓度 12. 暴露频率和持续时间 13. 是否有阴性或阳性对照 14. 每一个剂量组的实验动物数 15. 是否有可以用来评价研究的操作细节 ^g 16. 只针对吸入和重复剂量研究	8. 实验方法是否明确 9. 给予的剂量或浓度 10. 暴露频率和持续时间 11. 是否有阴性对照 12. 是否有阳性对照 13. 是否做了重复实验
第4组:实验结果的记录	17. 实验终点以及判断终点的方法是否明确 18. 所有试验终点研究结果的描述是否完整和透明 19. 数据的统计分析方法是否清晰的表达	14. 实验终点以及判断终点的方法是否明确 15. 所有实验终点研究结果的描述是否完整和透明 16. 数据的统计分析方法是否清晰的表达
第5组:实验设计和结果的可信性	20. 选择的实验设计对于获得化学品特异的实验数据是否恰当 21. 定量实验结果是否可靠	17. 选择的实验设计对于获得化学品特异的实验数据是否恰当 18. 定量实验结果是否可靠

注:^a受试物鉴别包括化学名称、CAS号、化学结构等,如果是混合物应明确各组成成分。^b来源指提供化学品的供应者(如进口商或生产者等)。^c例如吸入实验的颗粒物大小或刺激性实验的pH值等。^d所选实验动物种类,大鼠、小鼠等,并说明原因。^e人群实验要考虑人种,实验动物要考虑动物的级别,如SPF级动物。^f温度、湿度、食物、每笼动物数。^g例如饮食、介质、溶剂对受试物的稀释,经皮吸入实验皮肤的敏感性,总灌胃量。

2.3.3 ToxRTool 评价方法

实验数据和信息按照是否符合各项标准来回答“是”或者“否”,回答“是”得1分,“否”得0分,最后计算出总得分。表1中用红色标示的标准所涉及的信息,对于评价数据可靠性是必不可少的,称为“红色标准”(本文中用斜体字表示)。与至少一个红色标准不相符的研究数据将直接被分到Klimisch第3类,不考虑总得分。符合所有的红色标准的研究根据体内试验或体外实验计算不同的得分,按照总得分将毒性资料分成3类,具体见表2。

表2 ToxRTool 分类方法
Table 2 ToxRTool data classification method

分类	体内试验	体外试验
第1类	18~21分	15~18分
第2类	13~17分	11~14分
第3类	<13分或不满足所有的红色标准	<11分或不满足所有的红色标准

由此可见,ToxRTool所提出的评价标准更加详细,通过该工具所进行的可靠性评价更加清晰和透明,目前该工具已被用来辅助Klimisch分类,它对于化学品毒理学数据的可靠性评价更加规范和准确。

2.4 3种评价系统的对比

总的来说,上述3种评价系统的评价标准和评价方法存在着差别,但宗旨和原则是一致的,都是根据一定的评价标准对数据进行分类,选出可靠数据剔除不可靠数据,再对不同的类别进行不同的处理。评价毒理学数据时,3种评价系统各有利弊。Klimisch评级系统是REACH法规使用的的数据评价系统,而REACH法规是欧盟为建立统一的化学品监控管理体系而建立的一项法规,它的建立强化了欧盟在化学品领域的竞争优势,同时也对全球的化学品管理具有深远的意义。随着REACH法规的实施,许多非欧盟国家也开始制定类似的化学物质管理法规,可见REACH法规的影响力是十分明显的。因此,REACH法规推荐使用的Klimisch评级系统在毒理学数据评价中更具有权威性。但其在评价过程中也存在着一定的缺陷,主要表现在只是利用类别定义对数据进行评价,评价标准不明确,而这一缺陷恰恰是ToxRTool最主要的优点,评价标准明确具体,受主观因素影响少。美国EPA信息筛选系统的评价标准也比较明确,但是相比其他两个系统来说,评价过程较为繁琐,工作量大。所以,在建立我

国的毒理学数据评价系统中,建议将 Klimisch 评级系统和 ToxRTool 结合起来,在必要的情况下借鉴 EPA 信息筛选系统的评价模式,这样就能够做到客观、准确地对毒理学实验数据进行评价。

3 展望

化学品风险评价作为化学品管理的重要内容,一直以来在世界范围内受到广泛重视,并且近年来在技术上不断取得新的进展。化学品风险评估主要取决于实验数据的选择和有效性评价,目前,欧盟、USEPA 等国际组织使用几种毒理学数据可靠性评价工具,对数据进行可靠性评价,从而对化学品的危害及风险进行评估。本文论述的几种数据可靠性评价方法是目前国际上认可并且通用的,虽然有这些标准来评价数据,但评价工作中仍会出现不同评价者对于同一数据做出不同分类的现象^[15],即产生异质性,这可能与评价者的专业知识、对标准的理解和运用以及对可靠性的把握有关,这就需要不同的评价者之间相互探讨相互协商,最终确定数据的可靠性。

化学品安全管理逐渐趋于国际化,我国监管部门对其也更加重视,相继出台了《危险化学品安全管理条例》、《新化学物质环境管理办法》等法律法规,积极参加了一些国际行动和国际公约,并实施了农药登记、化学品首次进口和有毒化学品进出口登记、合格实验室认证等管理措施,取得了一定的成效。然而由于中国地域广阔,涉及化学品的企业多且分散,很难实施集中高效的管理,因此在法规的建立与完善、能力建设、技术支持体系的建立等诸多方面还有许多实际问题需要解决,以适应国际化的发展趋势。我国要想在化学品风险评估方面取得进一步的发展,当务之急就是要解决以下问题:建立和完善地方性法规和部门规章,加强各地区、各部门之间的有效配合;加快化学品测试合格实验室认证体系的建立和完善,实现实验数据的互认与共享;加强国际及国内各部门之间的信息交换和交流,有效地利用现有的数据和信息资源,为化学品管理服务。

参考文献

[1] 菅小东,周红,高映新. 国际化学品安全管理的现状及发展趋势

势[J]. 环境科学研究,2004,17(3):7-9.

- [2] EU. REACH (regulation for the registration, evaluation, authorisation and restriction of chemicals) [S]. 2007-06-01.
- [3] KLIMISCH H J, ANDREAE M, TILLMANN U. A systematic approach for evaluating the quality of experimental toxicological and ecotoxicological data[J]. Regul Toxicol Pharmacol, 1997, 25(1):1-5.
- [4] European Chemicals Agency. REACH "Guidance on Information requirements and Chemical Safety Assessment", Chapter R. 4 "Evaluation of available information" [EB/OL]. [2010-08-11]. http://guidance.echa.europa.eu/docs/guidance_document/information_requirements_en.htm.
- [5] WHO. Evaluation and use of epidemiological evidence for environmental risk assessment: guidance document [R]. Copenhagen: WHO, 2000.
- [6] WEED D. Weight of evidence: a review of concept and methods [J]. Risk Anal, 2005, 25(6):1545-1557.
- [7] MARKWIESE J T. Collection and evaluation of existing data: an ecological risk assessment perspective [J]. Qual Assur, 2003, 10(3-4):107-116.
- [8] Environmental Protection Agency. High Production Volume (HPV) Challenge Program [R]. EPA, 1999.
- [9] SCHNEIDER K, SCHWARZ M, BURKHOLDER I, et al. "ToxRTool", a new tool to assess the reliability of toxicological data [J]. Toxicol Lett, 2009, 189(2):138-144.
- [10] HEIDORN C J, RASMUSSEN K, HANSEN B G. IUCLID: an information management tool for existing chemicals and biocides [J]. J Chem Inf Comput Sci, 2003, 43(3):779-786.
- [11] UNEP. Screening Information Data Sets for High Production Volume Chemicals [R]. Paris: OECD, 2006.
- [12] U. S. Environmental Protection Agency. Determining the Adequacy of Existing Data, Guidance for the HPV Challenge Program, Draft dated 2/10/99 [EB/OL]. (2010-08-02) [2010-08-30]. <http://www.epa.gov/hpv/pubs/general/datadq23.htm>.
- [13] OECD (2005a) -Manual for Investigation of HPV Chemicals. Chapter 3-Data Evaluation. Organization for Economic Cooperation and Development [EB/OL]. [2010-10-14]. http://www.oecd.org/document/7/0,3343,en_2649_34379_1947463_1_1_1_1,00.html#Chapter_1.
- [14] European Centre for the Validation of Alternative Methods. ToxRTool-Toxicological data Reliability Assessment Tool [EB/OL]. [2010-11-08]. <http://ecvam.jrc.it/>.
- [15] HOBBS D A, WARNE M S, MARKICH S J. Evaluation of criteria used to assess the quality of aquatic toxicity data [J]. Integrat Environm Assess Manag, 2005, 1(3):174-180.