

# 食品安全风险评估应用系统综述的 原则和方法

Systematic review of applications for food  
safety risk assessments: principles and  
methods

国家食品安全风险评估专家委员会

二〇一七年十二月

## 目 录

第一章	总则.....	3
第二章	系统综述的应用条件和基本流程 .....	3
第三章	系统综述的实施 .....	5
第四章	附则.....	11
附 录		
附录 1	系统综述和 meta 分析研究方案优先报告的项目清单 .....	14
附录 2	食品安全风险评估系统综述文献检索筛选的参考流程图 .....	17
附录 3	偏倚风险问题 .....	18

## 第一章 总则

1. 为了对食品安全风险评估数据进行系统检索和分析，提高我国风险评估工作科学性和规范性，根据食品安全风险评估工作要求，参照欧洲食品安全局（EFSA）和美国国家毒理学计划（NTP）等组织推荐并使用的相关程序和指南，制定本文件。

2. 本文件规定了食品安全风险评估（主要是危害识别步骤）应用系统综述方法的原则和方法，适用于食品安全风险评估工作中动物试验数据和人类研究资料的系统综述，可为我国食品安全风险评估相关工作提供参考。

3. 本文件遵循文献系统综述的基本程序和风险评估工作的基本要求，尽可能与国际相关技术指南保持一致，并在保证科学性与规范性的前提下，兼顾可操作性，以保证食品安全风险评估工作的顺利实施。

4. 本文件所指的系统综述(systematic review, SR)是一种收集、评估现有研究证据的结构化程序。通过该程序可以在评估某一特定问题现有研究证据的基础上，做出科学决策。一般来说，高质量的系统综述结论被视为最高水平的危害识别证据。

## 第二章 系统综述的应用条件和基本流程

5. 在食品安全风险评估中，建议使用系统综述方法，为阐述食品中某物质的健康效应提供科学证据，尤其是当针对某一问题已经积累了足够多的研究证据且研究结论不一致时。

6. 在食品安全风险评估中应用系统综述时，应将当前的研究证

据分为人群研究资料和动物试验资料，分别进行资料分析和危害识别。

7. 在应用系统综述进行危害识别时，主要工作包括文献的检索和筛选、文献的质量评价、不同类型证据（人群资料、动物试验）的等级评价、所有证据的整合等，可参照如下步骤进行。

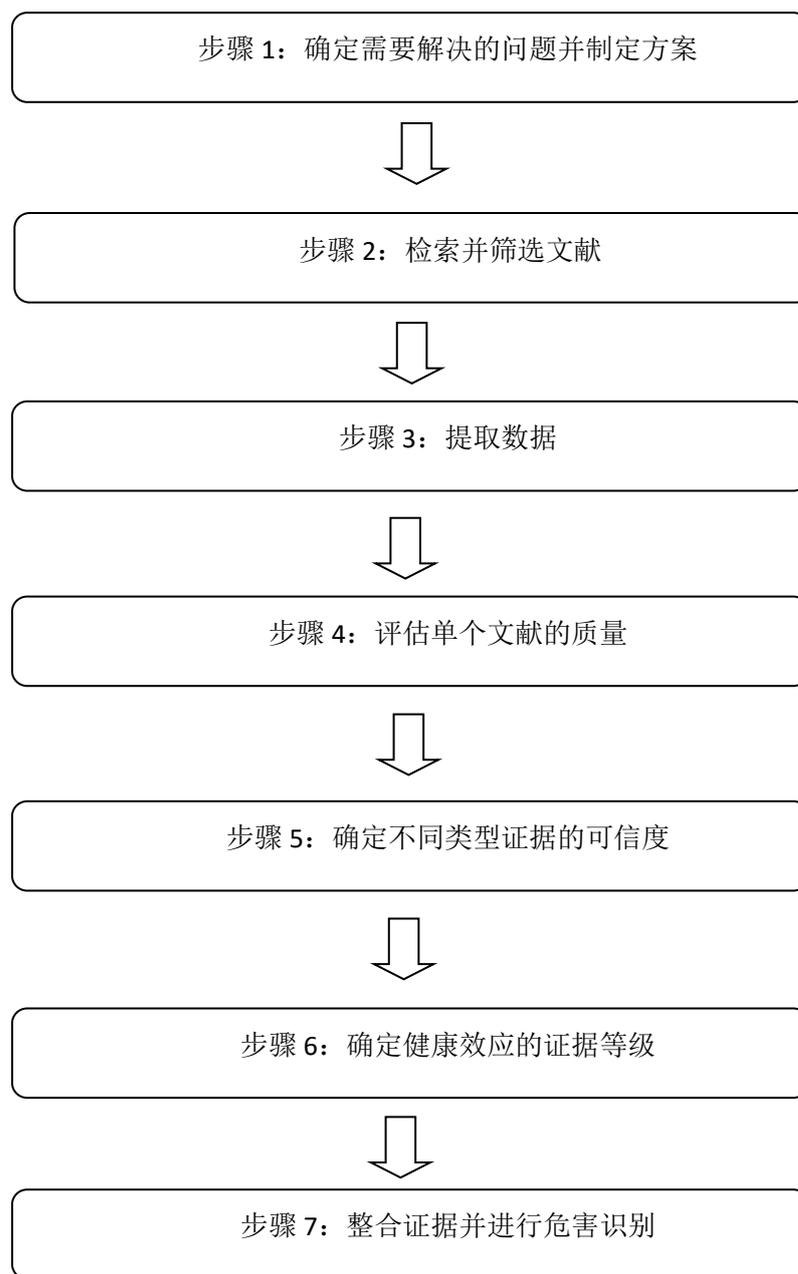


图 1 基于系统综述的食品安全危害识别流程

### 第三章 系统综述的实施

8. 第一步：确定需要解决的问题并制定方案

9. 任何风险评估（危害识别）都应以问题为导向，应依据食品安全风险评估的任务需求，确定系统综述需要解决的问题，如“化学物质 X 是否具有某系统毒性？”等。

10. 危害识别的问题可参照系统综述方法中的 PECO/PICO 原则来确定，即：目标群体（P）：指整个人群或者某个特定的人群分组，实验动物种群等。暴露因素（E）或干预因素（I）：指食品中某类物质的暴露，包括食品中的污染物、添加剂等待评估物质。比较组（C）：指用于对比的参照对象，例如人群或动物试验中的对照组、不同剂量水平的群体等。结局（O）：指暴露所产生的（不良）健康效应。

11. 由于动物实验数据中单一暴露产生的结局较多，在食品安全风险评估中运用系统综述进行危害识别时，尽可能明确需要评估的（不良）健康结局，必要时需要利用毒理学等专业知识确定反映健康结局的主要指标和次要指标，以及指标的优先顺序。

12. 制定工作方案时应依据系统综述确定的问题，参考国际上通用的系统综述和 meta 分析研究方案优先报告的项目清单进行（见附录 A 的 PRISMA-P）。为了增加透明度和获得其他专家的专业评议，工作方案可根据需要适当公开或在系统综述相关注册机构（如 PROSPERO）上登记。

13. 第二步：检索并筛选文献

14. 根据需要解决的危害识别问题，确定检索词及相应的限定条

件（如对检索日期、文献语种的限制），并在系统综述方案中注明。为了避免文献漏检，可通过增加相关的检索词来提高检索的广泛性和覆盖性。

15. 在文献检索时，应尽可能覆盖现有的各种类型数据库，包括但不限于 Pubmed, TOXNET, Scopus, Embase, Web of Science, Google Scholar, CNKI, 万方数据库等，学位论文、会议资料以及可获得的尚未公开的技术报告也可以作为重要文献，并在系统综述方案中注明各数据库检索时采用的完整检索策略。

16. 应根据 PECO/PICO 原则确定的结局类型、暴露种类、研究类型等因素来制定文献的纳入与排除标准。文献的检索、纳入、排除过程应当参考附件 B 建议流程图的架构，应详细记录文献纳入、排除的数量和流程以及文献排除的原因，同时强调有几位审查人员进行文献审查，并指出审查人员是否独立进行文献审查工作。

17. 建议使用类似 EndNote, NoteExpress 等文献管理软件对文献进行管理。

18. 第三步：提取数据

19. 在对符合纳入标准的单个文献资料进行数据提取时，应包括 PECO/PICO 原则确定的群体特征、暴露水平、结局以及相关影响因素等信息。应根据系统综述方案所确定的表格进行数据提取。

20. 针对某一特定结局，应按照人群资料和动物实验资料分别进行数据提取，形成人群资料和动物实验资料不同类型的证据集。

21. 第四步：评价文献的质量

22. 文献质量评价通常涉及可靠性、相关性和充分性三方面，其中可靠性反映文献的内在质量，是评价相关性和充分性的必要前提。原则上，应在完成可靠性和相关性评价后，首选总体质量较高的文献进入下一步程序。

23. 目前，国际上尚无一致公认的文献质量评价方法和程序，现有不同机构使用的质量评价工具存在一定差别。进行文献质量评价时，可根据文献类型和特定情形参照相关方法进行。例如，可以参照《食品安全风险评估毒理学数据可靠性评价指南》或美国 NTP 推荐的偏倚风险评价方法（见附录 C）进行文献可靠性评价。

#### 24. 第五步：确定证据的可信度

25. 证据的可信度是指证据能够证明某物质暴露与健康结局之间是否存在实际关联的强度。通常在汇总了所有数据或文献证据后，将人类研究证据和动物资料证据分开进行评价。

26. 证据的可信度评价通常分两步进行，先进行初始评价，然后根据影响可信度的因素清单进行分析，调整分级。在初始评价过程中，首先评价某类证据是否符合以下四个研究设计特征以及符合程度，包括：（1）对特定暴露物质是否可控；（2）暴露是否发生于结局出现之前；（3）结局是基于个体而非群体进行评价的；（4）研究中是否有比较组。不同设计的人群资料的证据可信度初始评价分级参考见表 1。各种类型研究设计所获得的证据初始分级应在此基础上根据其具备研究特征要点的数量进行调整。

表 1 不同设计的人群资料的可信度初始评价参考

研究设计	初始评价
人群 RCT 试验	高
人群非 RCT 试验	中
队列研究	中
病例对照	低
横断面研究	低
生态学研究	低
病例报告	极低

27. 动物实验资料通常经过严格设计，与四个研究设计特征的符合程度高，其证据可信度的初始评价通常较高。

28. 对初始评价结果进行调整时，通常需要考虑的影响因素包括偏倚风险程度、所有证据间的一致程度、研究结果的精确程度等。可参照表 2 和表 3 相关内容，降低或增加初始评价过程中获得的证据可信度级别。

29. 证据的可信度一般分为高、中、低、极低 4 级，具体说明见表 4。某类证据（人群研究或动物实验）的可信度结论取决于该类证据的最高可信度。

表 2 降低证据可信度级别的影响因素

项目	说明
偏倚风险	指大多数的研究均存在偏倚风险
无法解释的非一致性	指不同研究结果的方向和程度彼此差异很大
间接性	指对研究结局的非直接测量，例如研究设计的人群、暴露、结局的测量与目标人群不同。
不精确	指研究结果缺少确定性（如效应的可信区间太宽）

发表偏倚 指发表文献的偏倚，例如仅有少量小样本阳性结果发表

**表 3 增加证据可信度级别的影响因素**

项目	说明
效应值大	指效应足够大以至于难以用潜在混杂偏倚解释
存在剂量反应关系	指暴露水平与结局之间存在剂量反应关系
合理的残余混杂	指研究中未进行校正的偏倚可能导致实际效应被低估，则会增加研究结果的可信度。
结果一致性	指在不同种族人群或物种间得到相同或类似的结论
其他因素	

**表 4 证据可信度分级**

可信度分级	定义
高	非常确信暴露与结局关联的证据。关联有很大可能性反映了真实的效应
中	中等确信暴露与结局关联的证据。关联可能反映了真实的效应
低	对暴露与结局关联的证据信心有限。关联可能与真实的效应不同
极低	对暴露与结局关联基本不可信。关联有很大可能性与真实的效应不同

30. 如果所有证据的可信度被评价为“极低”，说明目前尚无可信的证据来反映暴露与健康效应的关联，则无需进入下一步程序。只有证据的可信度为“中”以上时，才有必要进入下一步程序。

31. 如果所获得的数据满足 meta 分析条件，即人群或动物研究对象一致，评估指标相同，并且有足够的研究数据（一般多于 5 个研究或数据），应按照 meta 分析流程，估计特定暴露导致健康效应的综合值。

32. 第六步：确定健康效应的证据级别

33. 将第五步中证据的可信度评价结果转化为对暴露的健康效

应的证据等级。原则上，当证据指向暴露与结局有关联时，证据可信度的高、中、低 3 个级别可直接转化为相应的健康效应证据等级（见下表）。当证据指向暴露与结局无关联时，高可信度的证据可转化为“无健康效应”证据，中、低可信度的证据直接转化为“无充足证据表明无健康效应”。

34. 需要说明的是，由于证明阴性结果相对困难，“无健康效应”的结论必须在有充足的高水平证据的基础上才能够得出。

**表 5 健康效应的证据水平**

等级	定义
高水平证据	新的研究不可能改变现有资料得出的结论；现有结论是根据一组已经合理排除偶然性、偏倚和其他因素的资料得出。
中等水平证据	新的研究可能影响现有资料对结果的解释，现有结论是根据一组未能合理排除偶然性、偏倚和其他因素的资料得出。
低水平证据	现有资料存在一些局限性，诸如研究设计的灵敏度或相关性不足，不能做出判断，或者没有可用的资料。

### 35. 第七步：危害识别

36. 危害识别结论可分为四类，从高到低分别为：确定对人类有害、很可能对人类有害、怀疑对人类有害和无法确定是否对人类有害。

37. 进行待评估物质危害识别时，可参照图 2 的综合判定矩阵，将上述人群资料和动物资料的健康效应证据进行整合，通常优先使用最高等级的证据资料。

38. 在做出最终结论时，需要充分考虑其他的相关研究。这些相关研究证据（包括但不限于体外试验证据或者机制方面的证据）可能

对上述危害识别结果产生影响（如图 2 中箭头部分），例如：当机制研究的证据在生物学方面强烈支持暴露与健康效应的关联性，可考虑将上述危害识别级别增加 1 级。

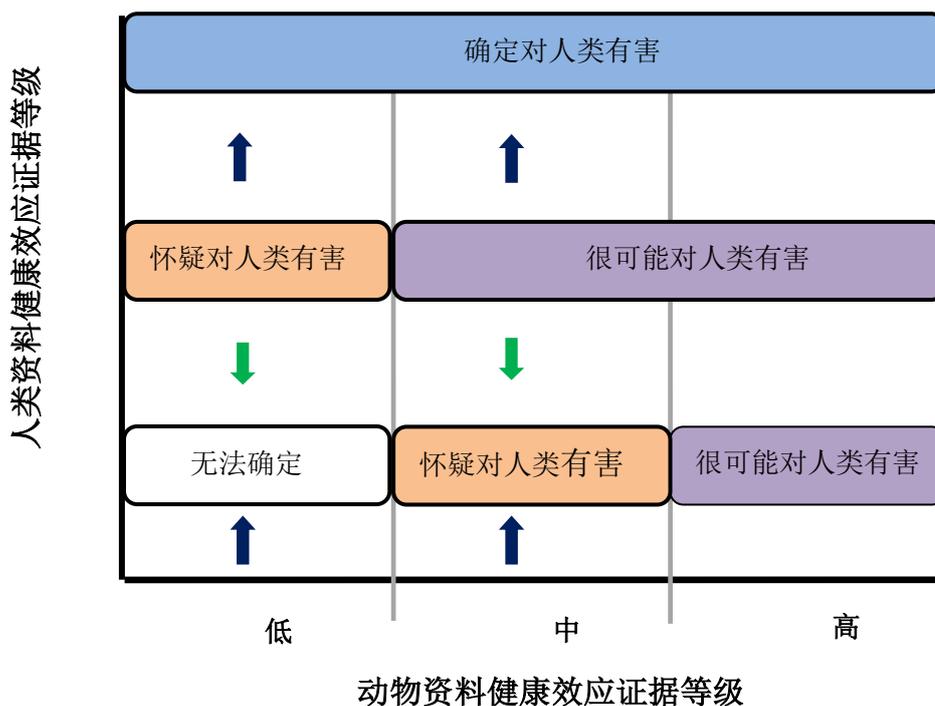


图 2 整合人群资料和动物资料证据等级的危害识别矩阵

#### 第四章 附则

39. 为避免专业局限和学术偏向，应组建包括毒理学、流行病学与统计学、基础医学等领域的专家组成的专家小组开展基于系统综述的危害识别。建议由两名专业人员分别独立进行文献筛选、质量评价、证据分级等工作，完成后再进行互相核对确认。当结论不一致时，应由团队中的资深成员确认或通过专家小组讨论，达成一致后确认最终结果。整个实施过程应进行详细记录并存档。

40. 在食品安全风险评估的危害识别中，同一个健康结局（如免疫毒性）的资料通常涉及不同层级的指标（如免疫器官变化、淋巴细胞亚型和计数、血液中白介素含量等），在进行证据可信度分级时，应对不同层级的指标进行合并，形成针对该健康结局的最终证据可信度分级。此时需要慎重考虑不同层级指标在该健康结局中的重要程度。

41. 采用系统综述方法进行危害识别，应遵循风险评估的基本原则。对已完成的系统综述（特别是结论无法确定对人类是否有害时），需要定期重新评价新出现的证据，视需要重新对证据的可信度、证据等级进行分析，以便及时调整危害识别的结论。

42. 本文件引用的其他相关术语和定义包括：

（1）**Meta 分析（meta analysis）**：系统综述中定量合并效应估计的数据分析方法，可以通过合并多个研究结果，定量估计效应的大小，是一种数据综合分析方法。

（2）**PRISMA 声明（The Preferred Reporting Items for Systematic reviews and Meta-Analyses）**：即系统综述与 meta 分析优先报告条目，是 2009 年由多国专家联合在多个重要杂志上发布的用于规范系统综述和 meta 分析报告的质量，并于 2020 年修订。该声明在世界范围内得到了广泛的应用，并衍生出相应的扩展版本和细化版本用于规范各类专题系统综述的报告质量。

（3）**GRADE（The Grading of Recommendations Assessment, Development and Evaluation）**：推荐分级的评价、制定和评估，是在世

界范围内得到广泛关注的证据质量分级和推荐强度系统。它由 GRADE 工作组 2004 年正式推出的，并在 2011 年进行了重大修订，它采用易于理解的方式评价证据质量和推荐等级，在系统评价、卫生技术评估以及指南的制定上得到了广泛的应用。

## 附录

附录 1 系统综述和 meta 分析研究方案优先报告的项目清单  
(PRISMA-P, 2015 年版)

项目	项目编号	项目说明
管理信息		
题目		
识别信息	1a	题目要能显示本报告为系统综述研究方案
更新	1b	如果该方案为以前系统综述的更新，应注明
登记	2	如果系统综述方案已在某些机构登记应注明
作者		
联系方式	3a	应提供方案设计者的联系方式、机构名称和全部作者清单
贡献	3b	应注明各作者的贡献及方案负责人
修订	4	如果研究方案是以前已发表方案的修订，应注明
支持		
资源	5a	应注明资金资助来源或其他支持来源
资助者	5b	提供资助者的姓名
资助者的角色	5c	描述资助者在开本研究方案中的角色
前言		
合理性	6	在介绍目前已知结果基础上描述系统综述的合理性
目标	7	参考PICO/PECO原则描述系统综述的问题
方法		

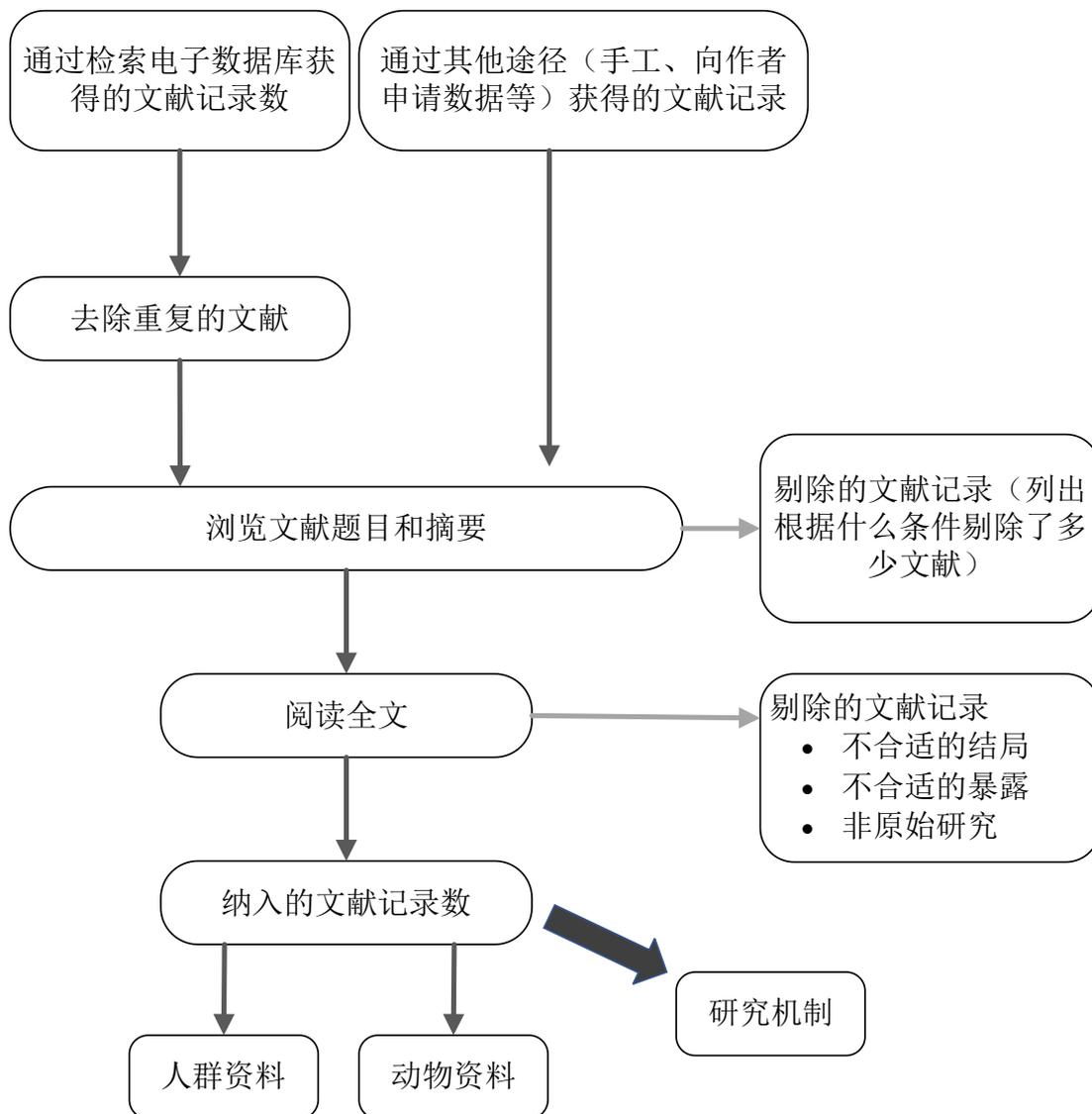
---

合格标准	8	定义研究的特征作为选择合格研究进入综述的标准
信息来源	9	描述用于检索的各种信息来源（例如电子数据库，试验登记以及其他信息来源）
检索策略	10	至少展示一个用于电子数据库检索的策略，以确保其可以被重复检索
研究记录		
数据管理	11a	描述在整个综述过程中数据管理的方法
选择过程	11b	描述纳入综述的研究的选择过程
数据搜集过程	11c	描述提取数据的方法（例如提取数据表格以及数据提取结果的确认方法等）
数据项目	12	列出所有纳入综述的数据变量的清单和定义
结局和优先顺序	13	列出所有检索结果所涉及的结局，确定主要结局和次要结局，以及结局的优先顺序。
单个研究的偏倚风险	14	描述评估单个研究偏倚风险的方法，并注明如何使用这些信息进入数据综合分析
数据综合分析	15a	描述进入数据综合分析的研究纳入标准
	15b	如果数据足够进行综合分析，应描述计划中的综合分析方法包括进行一致性检验的方法
	15c	描述进一步分析的方法，例如敏感性分析，分组分析，meta回归分析等

	15d	如果数据不适合进行综合分析，描述进行数据汇总的方法
meta-偏倚	16	定义任何对meta偏倚评价的方法，例如发表偏倚、选择报告偏倚等。
综合证据的可信度	17	描述如何评估综合证据的可信度，例如GRADE系统评价方法

---

## 附录 2 食品安全风险评估系统综述文献检索筛选的参考流程图



## 附录3 偏倚风险问题

偏倚种类及问题	适用的研究设计
<b>选择偏倚</b>	
暴露水平或实验剂量分组是否随机化	ExA, HCT
研究对象是否进行了恰当的隐匿分组?	ExA, HCT
比较组是否合适?	Coh, CaC, CrS
<b>混杂偏倚</b>	
研究设计或分析是否考虑了重要的混杂变量或效应修饰变量	所有设计
研究者是否调整或控制了其他可能影响结果的暴露因素	所有设计
<b>实施偏倚</b>	
所有试验组的实验条件是否完全一致?	ExA
研究者是否严格遵循研究方案?	所有设计
研究者和研究对象是否对于研究分组处于盲法状态?	ExA, HCT
<b>自然损耗/排除偏倚</b>	
结局数据是否完整没有或很少有自然损耗或排除	ExA, HCT, Coh, CaC, CrS
<b>测量偏倚</b>	
是否结局的测量者对于实验分组或暴露水平处于盲法状态?	所有设计
是否混杂变量的评价在各研究组都使用了有效可靠的测量方法?	所有设计
我们对暴露的测量是否可信?	所有设计
我们对结局的测量是否可信?	所有设计
<b>选择报告偏倚</b>	
是否所有测量的结局都进行了报告?	所有设计
<b>其他</b>	

是否有其他影响内部有效性的潜在风险？（例如：统计分析 任何对于特定研究  
方法是否适当） 设计需要考虑的其  
他因素

---

CaC: 病例对照研究; CaS: 病例系列研究; Coh: 队列研究; CrS:  
横断面研究; ExA: 动物实验研究; HCT: 人类试验研究