

# 行业指南：基于证据的用于科学评估健康 声明的审查系统(终稿)

2009 年 1 月

包含非约束性建议

欲索取本文件，请与以下部门联系：

*Office of Nutrition, Labeling and Dietary Supplements  
Division of Nutrition Programs and Labeling HFS-830  
Center for Food Safety and Applied Nutrition  
Food and Drug Administration  
5100 Paint Branch Parkway  
College Park, MD 20740  
(电话) 301-436-1450  
<http://www.cfsan.fda.gov/guidance.html>*

您可以随时以书面或电邮形式对本指南提出意见。请把有关指南的书面意见发送至：Division of Dockets Management (HFA-305), Food and Drug Administration, 5630 Fishers Lane, rm. 1061, Rockville, MD 20852。电子邮件可发送至：  
<http://www.regulations.gov>。所有意见都应当注明文件号码，该文件号码载于《联邦公告》( *Federal Register* ) 中的发行通知。

美国卫生与公众服务部  
食品与药物管理局  
食品安全与应用营养中心  
2009 年 1 月

---

包含非约束性建议

目录

1. [引言](#)
2. [背景信息](#)
3. [基于证据的用于科学评估健康声明的审查系统](#)
4. [参考资料](#)

---

包含非约束性建议

## 行业指南<sup>[1]</sup>

### 基于证据的用于科学评估健康声明的审查系统

本指南代表食品与药物管理局 (FDA) 当前对该话题的意见。它并不形成或赋予任何人以任何权利，也不会对 FDA 或公众产生约束力。您可以采用替代方法，但该方法必须满足有关法律和法规的要求。如果您希望商讨一种替代方法，请与负责实施本指南的 FDA 工作人员联系。如果您无法确定应该与哪位 FDA 工作人员联系，请拨打本指南封面上列出的有关电话号码。

## 一、引言

本文件为企业提供指南。它代表了目前 FDA 对以下问题的意见：1) 用于评估健康声明科学证据的程序，2) 《联邦食品、药物与化妆品法》(简称《法案》) 第 403(r)(3) 条 (《美国法典》第 21 卷第 343(r)(3) 条) 和《美国联邦法规》第 21 卷第 101.14(c) 条中重大科学共识 (SSA) 标准的含义，以及 3) 用来支持合格健康声明 (QHC) 的可靠的科学依据。

本指南所描述的是基于证据的审查系统。FDA 会使用该系统来评估公开的、用于 SSA 健康声明或合格健康声明的科学证据，该类声明涉及某种物质与某种疾病或健康问题之间的关系。<sup>(2)</sup> 本指南说明了 FDA 目前对其应当使用的科学审查方法的意见，并为申请健康声明者提供指导。<sup>(3)</sup>

本指南所涉及的具体题目有：(1) 认定评估物质与疾病关系的研究；(2) 确定疾病风险替代终点指标 (surrogate endpoints)；(3) 评估人体研究，以确定是否能

从中得出有关物质与疾病关系的结论；(4) 评估每一项人体研究所用方法的质量，该研究可能得出物质与疾病关系的科学结论；(5) 评估科学证据的完整性，(6) 评估重大科学共识，(7) 合格健康声明的语言特征，以及(8) 重新评估现有的 SSA 或合格健康声明。

FDA 的指南文件（包括本指南）并不规定在法律上必须履行的责任。相反，指南文件只是描述 FDA 当前对某一问题的意见，只应视作建议，除非引用了具体规章或法律要求。在 FDA 指南文件中，“应当”一词是指提出建议，而不是强制性规定。

## 二、背景信息

《1990 年营养标签和教育法》（NLEA）（出版物 L. 101-553）旨在促进为消费者提供关于食品的更具有科学性的信息。除其他条款外，NLEA 指导 FDA 发布规章，要求在食品（包括食品添加剂）标签中包含描述物质与疾病关系的声明（“健康声明”），但此等声明须经 FDA 事先审查和授权。<sup>(4)</sup>就这些健康声明——即关于物质与疾病关系的声明——而言，FDA 通过规章把“物质”定义为一种具体的食品或食品成份（《美国联邦法规》第 21 卷第 101.14(a)(2) 条）。经授权的健康声明既可用于常规食品又可用于食品添加剂，前提是产品中的物质和产品本身符合授权规章中的有关标准。健康声明针对公众或特定群体（例如老年人），旨在帮助消费者保持健康的饮食习惯。

在评估一项健康声明的申请时，FDA 考虑该声明内容中支持其关系的证据是否符合 SSA 标准。该标准源于《美国法典》第 21 卷第 343 (r)(3)(B)(i) 条，它规定 FDA 可授权用于常规食品的健康声明，前提是 FDA 认定，“基于所有的公开信息（包括设计良好、符合公认的科学程序和原则的研究所提供的证据），在经过受过科学训练、拥有科学经验、有资格评估此类声明的专家中存在着相当普遍的共识，且该声明受到此类证据的支持。”法律规定，常规食品健康声明一律采用该科学标准；在法规中 FDA 对食品添加剂健康声明也采用了同样的标准。请参见《美国联邦法规》第 21 卷第 101.14(c) 条。

合格健康声明起源于上诉法院对 *Pearson v. Shalala (Pearson)* 一案的裁决。在该案件中，对关于食品添加剂标签中的四种特定的物质与疾病关系的健康声明，FDA 未予授权，原告就此提出异议。虽然地区法院裁定 FDA 胜诉（法院案例 14 F. Supp. 2d 10 (D.D.C. 1998)），但哥伦比亚特区美国巡回上诉法院推翻了下级法院的裁决（法院案例 164 F.3d 650 (D.C. Cir.1999)）。上诉法院认为，第一修正案不准许 FDA 拒绝它认为可能会误导公众的健康声明，除非 FDA 有理由认定通过免责声明仍无法消除可能存在的误导。上诉法院还认为，《行政诉讼法》（APA）要求 FDA 澄清授权健康声明时采用的“明确的科学共识”（SSA）标准。

1999 年 12 月 22 日，FDA 宣布发布其《行业指南：常规食品和食品添加剂健康声明审查中明确的科学共识》（《联邦纪事》第 64 卷 17494 页）。发布该指南文件旨在回应上诉法院在 *Pearson* 一案中的第二项意见，澄清 FDA 对 SSA 标准的解释。

2002 年 12 月 20 日，FDA 宣布它要扩大执行 *Pearson* 案裁决的范围，把常规食品健康声明也包括在内（《联邦纪事》第 67 卷 78002 页）。FDA 认识到有必要建立一个合格健康声明的科学框架，于是便成立了“改善消费者营养健康信息”

（Consumer Health Information for Better Nutrition）工作组。该工作组认识到，一旦公众在购买时能够从常规食品以及食品添加剂标签中获得和使用更多更好的信息，就会对公共健康产生显著的益处。这是因为这些信息不仅包含了价格、便利程度和口味，而且还延伸到了以科学为基础的健康因素。消费者如果掌握了更多的关于食品和食品添加剂对健康影响的科学信息，就能够降低诸多慢性病风险，对其长期健康产生有益的影响。

为了最大限度地发挥 FDA 声明审查程序对增进公众健康的益处，工作组在《最终报告》中<sup>⑤</sup>提出一种程序，用来对所有完整的申请以个案为基础，按照多种因素进行排序。排序的因素中包括作为申请对象的食物或食品添加剂是否可能对一种严重的或危及生命的疾病产生显著影响；证据的力度；是否提供了显示该声明没有误导性的消费者调查；作为声明对象的物质是否经过 FDA 安全审查（即是否属于经授权食品添加剂，是否被确认属于“公认安全”（GRAS）物质，被列入清单，或是否就 GRAS 通知收到“不反对”信函）；作为声明对象的物质是否已经得到恰当的定性，

以便评估现有研究的相关性；疾病是否按照由一些公认的有资质的专家制定的公认标准来定义和评估；证据或声明是否已先期得到一些公认的有资质的专家的审查。

作为工作组最终报告的部分内容，FDA 建立了一个基于证据的过渡性审查系统，并打算把该系统用于评估作为合格健康声明对象的物质与疾病的关系。在审议 1999 年 12 月 22 日 SSA 指南文件和 2003 年工作组报告时，FDA 明确认识到，SSA 健康声明与合格健康声明的科学审查程序之要素非常相似。由于这两类声明的科学审查程序相似，FDA 会用本指南中阐述的方法对为 SSA 健康声明或合格健康声明提交的申请中所包含的科学根据进行审查。本指南中阐述的基于证据的审查系统将有助于 FDA 认定提交的科学根据是否符合 SSA 标准；如果不符合，则认定该证据是否支持合格健康声明。除了科学审查外，健康声明还需要经过法规审查。符合 SSA 标准的健康声明通过在《联邦纪事》中发布的最终决定或过渡性最终决定来获得授权。至于有可靠证据支持的合格健康声明，FDA 会发出一封信函，表明其考虑执法自行裁量 ( enforcement discretion ) 的意向。

针对上诉法院在 *Pearson* 一案中的第二项裁决，FDA 颁布了《行业指南：常规食品和食品添加剂健康声明审查中明确的科学共识》(《联邦纪事》第 64 卷 17494 页)，以澄清 FDA 对 SSA 标准的解释。虽然本指南取代了该文件，但 FDA 认为本指南依然符合法院的裁决，1999 年指南中阐述的 SSA 基本原则没有改变。关于 SSA 的决定仍然要求 FDA 作出最佳判断，以确定有资质的专家们是否有可能同意用来支持拟议健康声明中的物质与疾病关系的科学证据。事实上，本指南中对 SSA 的许多说明与 1999 年指南中的语言完全相同。本指南反映的是 FDA 在审查科学证据的方法上进一步科学化，并未改变它对 SSA 内容的理解。

### 三、基于证据的用于科学评估健康声明的审查系统

#### A、何为基于证据的审查系统？

基于证据的审查系统指系统性地科学评估用来支持某项声明的证据的力度。对于健康声明而言，该系统评估的是某项拟议声明中用来支持物质与疾病关系的科学证据的力度。该评估程序涉及一系列步骤，包括评估科学研究和其他数据、排除无法从中得出物质与疾病关系结论的研究与数据、对余下的研究进行研究方法质量鉴定并

评估科学证据的完整性，其具体做法是审查研究类型、研究方法的质量、支持和反对该声明的证据的数量（考虑到不同类型研究的数目和研究样本规模）、与美国人口或目标群体的相关性、支持拟议声明的研究结果的可复制性、以及证据的总体一致性。在审查科学根据的完整性之后，FDA 确定是否存在支持一项经授权健康声明的明确的科学共识，或是否存在支持一项合格健康声明的可靠证据。

## B、认定评估物质与疾病关系的研究

FDA 审查涉及物质与疾病关系的公开数据和书面信息，包括健康声明申请中必须提及的各项研究（《美国联邦法规》第 21 卷第 101.70 条）。通过文献搜索，FDA 确定与拟议健康声明相关的其他研究。在审查物质与疾病关系证据的可信度之前，FDA 把关于人体研究的个别相关文章与其他类型的数据和信息区别开来。FDA 会把审查的重点放在报告人工干预和观察研究的文章，因为只有此类研究才能够提供有关物质与人类疾病关系的证据，从而得出科学结论。其次，在审查科学根据的过程中，FDA 还将考虑若干基本问题：

- *那些研究是否具体说明并测量了该声明涉及的物质？*研究必须认定可以测量的物质。“物质”的定义是一种具体的食品或食品成份，无论它是常规食品还是食品添加剂。（《美国联邦法规》第 21 卷第 101.14(a)(2)条）。例如，食品成份可以是一种营养物或食品配料。<sup>(6)</sup>如果在该物质作为常规食品成份使用时减少了食物【中其他食品成份（译者加注）】的摄入水平，则该物质就是一种营养物，就必须包括在“营养事实”标签中（《美国联邦法规》第 21 卷第 101.14(b)(2)条）。如果该物质不会减少其他食品成份的摄入水平，则该物质必须有助于口感、香味、营养价值，<sup>(7)</sup>或产生《美国联邦法规》第 21 卷第 170.3(o)条中列举的对食品的某种技术影响，而且其使用必须安全与合法，以证明该声明的合理性（《美国联邦法规》第 21 卷第 101.14(b)(3)条）。

- *那些研究是否适当指明并测量了该声明所涉及的具体疾病或健康问题？*“疾病或健康问题”的定义是对人体的器官、组成部分、结构或系统造成损害，使之不能发挥正常功能（例如心血管疾病）或可导致此类功能失常的健康状态（例如高血压）。（《美国联邦法规》第 21 卷第 101.14(a)(5)条）。研究可以通过测量发病率、相关

死亡率或经过验证可预测具体疾病风险的替代终点指标指标来确认一种具体的、可测量的疾病或健康问题。

例如，癌症包含 100 多种不同的疾病，每一种疾病都表现为异常细胞的过度增长和扩散（美国癌症学会，2004）。按照受影响的器官和组织部位，癌症被分为不同的疾病类别（国家癌症研究所）。处于不同器官部位的癌症有不同的风险因素、治疗方式和死亡风险（美国癌症学会，2004）。遗传和环境（包括食品）因素会影响患不同类别的癌症的风险。风险因素可包括每一类癌症的家族病史、吸烟、饮酒、超重或肥胖、接触紫外线或电离辐射、接触致癌化学品、食品因素等。每一类癌症的病因、风险因素、诊断和治疗方法各不相同（Hord et al., 2007; Milner et al., 2006）。由于每一类癌症基于器官位置、风险因素、治疗选择和死亡风险都是一种独特的疾病，FDA当前的方针是分别评估健康声明或合格健康声明申请中的每一种癌症，以确定就该类癌症是否存在有支持物质与疾病关系的科学根据，这种癌症构成《美国联邦法规》第 21 卷第 101.14(a)(5)条所规定的疾病。FDA遵循此方针发布了数封执法自行裁量信函，包括[2005 年 6 月 30 日发布的《绿茶和癌症》\\*](#)、[2005 年 11 月 8 日发布的《西红柿/番茄红素和各类癌症》\\*](#)、[2005 年 11 月 12 日发布的《钙和各类癌症》\\*](#)、以及一份载于《联邦纪事》的通知，标题是《健康声明与合格健康声明；膳食脂肪与癌症、大豆蛋白质与冠心病、抗氧化剂维生素与某些癌症以及硒和某些癌症；重新评估》（《联邦纪事》第 72 卷 72738 页, 2007 年 12 月 21 日）。

在考虑了这些关键性问题后，FDA 把各项研究分类。

## 干预研究

在干预研究中，干预组的被试通常得到常规食品或食品添加剂形式的相关物质（食品或食品成份）。应当对该物质的质量和数量进行控制。在随机对照试验中，被试被随机编组。个体被试之间的情况可能并不相同，但在随机抽样后，干预组和对照组的情况就应当相似了。随机对照试验是对物质与疾病之间因果关系的最佳评估，因为已知的干扰因素（即可能对疾病风险产生影响的其他因素）得到了控制。通过把被试按照随机方式分配至干预组和对照组，此类研究避免了选择偏差——即最有可能产生有益结果（无论是否干预）的被试被优先选入干预组。研究具有“盲性”，

即被试不知道自己是否接受干预，因而避免了可能的偏差。还有一种称为“双盲”的研究，即被试和评估结果的研究人员都不知道谁在干预组，谁在对照组。通过控制试验环境——包括有关物质的数量和成份——以及所有其他饮食因素，此类研究还能够最大限度地减少变量或干扰因素对结果的影响。<sup>(8)</sup>因此，随机对照干预研究提供最有力证据，能够说明物质和疾病之间是否存在关系(Greer et al., 2000)。

另外，此类研究还能够为相关干预与结果之间的因果关系提供令人信服的证据 ( Kraemer et al., 2005 at 113 )。但是，随机分组可能导致对照组和治疗组之间被试特征分布不均匀 ( 例如，基线年龄或血液[血清或血浆]LDL 胆固醇水平显著不同 )。如果基线值显著不同，那么就很难确定研究结束时的差异是源于干预还是源于研究开始时业已存在的差别。当相关物质作为补充剂使用时，应当给对照组提供安慰剂。当物质是食品时，可能无法提供安慰剂，因此这种研究中的被试可能不具有盲性。尽管在这种情况下研究不具有盲性，但仍然需要有一个对照组才能从研究中得出结论。

随机对照试验通常采用平行设计或交叉设计。平行设计研究涉及测试组和对照组两组被试，两组同步分别接受有关物质或成为对照。在交叉设计中，在指定的阶段结束之后，干预组和对照组的所有被试互相对换。

虽然干预研究是确定因果关系的最可靠的研究类别，但把对特定群体的研究结果推广至不同的群体可能不具有科学性。例如，如果研究表明摄取某种物质与降低青少年糖尿病风险相关，该研究结果不应用来推论成年人的糖尿病风险。

## 观察研究

观察研究测量物质与疾病之间的关联。观察研究缺乏干预研究所具有的环境控制。观察研究最能够反映自由生活<sup>(9)</sup>的群体，有可能建立物质与疾病之间的关联。与干预性研究相反，观察研究不能确定观察到的某种关系是反映物质导致疾病风险降低的关系还是一种巧合 ( Sempos et al., 1999 )。由于在研究开始时没有基于各种疾病风险因素对被试进行随机分组，因此需要收集和调整疾病风险的已知干扰因素，以便最大限度地降低偏差。例如，应当收集关于每一位被试风险因素的信息 ( 例如年龄、种族、体重和吸烟情况 ) 并用于调整数据，以便对物质与疾病关系作出准确测



量。需要调整的风险因素针对研究的每一种疾病确定。例如，心血管疾病的风险与随着年龄增大；因此，需要对年龄作出调整，以便排除可能的干扰。

要确定作为声明对象的物质是否得到适当测量时，有必要重点评价食品摄入估算方法。许多观察研究依赖被试自己报告的饮食情况（例如饮食记录、24小时回忆、饮食历史和食品频率问卷），这些报告只是对摄入食品的大致估计（国家研究理事会，1989）。饮食记录所基于的前提是食品重量能够准确反映食品摄入量。被试把他们食用的食品过称，记录其重量。24小时回忆方法要求被试描述在过去24小时内他们食用的食品以及每种食品的份量。食品历史方法利用问卷或访谈估计被试在某一阶段内的典型饮食。食品频率问卷是在大型食品和健康观察研究中最常用的食品估算工具。经验证的食品频率问卷在评估“普通”食品摄入量方面的可靠性高于饮食记录方法或24小时回忆方法(Subar et al., 2001)。该问卷让参加者报告在某个指定的阶段中食用某些食品的频率和份量，报告的食品是从一个食品清单中选择的。上述的食品估算方法的问题之一是对某些食品的自我报告可能会出现偏差。例如，超重的人倾向于少报其食品份量（Flegal et al., 1999），因此往往低估其实际摄入的物质。如果某种物质有可靠的摄入生物标记<sup>(10)</sup>，则常常对这些生物标记进行测量，而不采用自我报告摄入量的方法。

观察研究可具有前瞻性或追溯性。此类研究可能出现不同形式的偏差（信息与选择）。<sup>(11)</sup>在前瞻性研究中，研究人员招收被试，在疾病发生之前观察他们。前瞻性观察研究对发病率和物质接触状况进行比较。在追溯性研究中，研究人员在疾病发生后查看被试的病历并/或对被试进行访谈。追溯性研究特别容易出现测算错误和回忆偏差，因为它们依赖被试对过去摄取的食物回忆。由于观察研究控制变量的能力有限，就常常会受到干扰因素的影响，例如综合物质与疾病的交互作用。

设计得当的观察研究能够提供有用的信息，发现可能的关联，以便通过干预研究检测（Kraemer et al., 2005 at 107）。与干预研究相反，即使设计最好的观察研究也不能建立干预与结果之间的因果关系（Kraemer et al., 2005 at 114）。而如上所述，干预研究则能够测试是否有证据表明某物质与降低疾病风险之间存在因果关系。在某些情况下，可从中得出科学结论的观察研究能够用于支持 SSA 或合格健康声明中

的物质与疾病关系。如下文所述，每一种观察研究设计都有其长处和缺点 ( Sempos et al., 1999 )。

*群组追踪研究 ( cohort studies )* 属于前瞻性研究，此类研究比较摄入与未摄入某种作为声明对象的物质的被试之间的发病率。由于物质的摄入先于发病，这种研究设计可以保证被试对于该物质的摄入不是由于生病而导致的。群组追踪研究能够生成相对风险估计 ( Szklo and Nieto, 2000 )。<sup>(12)</sup> 群组追踪研究被认为是最可靠的观察研究设计 ( Greer et al., 2000 )。

在*病例对照研究 ( case-control studies )* 中，对有病的被试 ( 病例组 ) 与没有病的被试 ( 对照组 ) 进行比较。<sup>(13)</sup> 对病例组和对照组都采用食品估算方法对以往的物质摄入量进行估算。此类追溯性研究常常问及在疾病确诊前至少一年的食品摄入状况，因此很难获得准确的摄入量估算。另外，一项关键的假设是食品摄入没有因疾病的进程或对病情的了解而改变。因此，病例对照研究的设计并不控制因疾病引起的或作为对疾病反应的摄入量变化。病例对照研究能够提出发病机率，对患病可能作出相对风险估计 ( Szklo and Nieto, 2000 )。<sup>(14)</sup> 病例对照研究的可靠性被认为低于群组追踪研究 ( Greer et al., 2000 )。

*嵌套式病例对照研究 ( nested-case control studies )* 或*病例群组追踪研究 ( case-cohort studies )* 使用从预先定义的群体中挑选的被试，例如一个参与正在进行中的群组追踪研究的人群。病例是群体中被诊断患病 ( 例如肺癌 ) 的被试。在嵌套式病例对照研究中，每确诊一个病例 ( 例如肺癌 )，就从有风险的群体中选出一个被试作为对照。在病例组研究中，对照组从基线群体中随机选择 ( Szklo and Nieto, 2000 )。此类研究可以计算相对风险或发病概率。嵌套式病例对照研究或病例群组追踪研究的可靠性被认为低于群组追踪研究，但高于病例对照研究。

*横向研究 ( cross-sectional studies )* 通常收集在某一时间点上患病者和无病者摄入食品的信息。<sup>(15)</sup> 此类研究可用于确定可能的相关性 ( 如确定对某种物质的摄入量与发病率之间的相关系数<sup>(16)</sup> )，并为以后的前瞻性研究提供基线信息 ( Kraemer et al., 2005 at 99-100 )。但是，由于同期测量食品摄入量和疾病状况，就无法确定物质的摄入是影响疾病风险的因素还是患病的结果。横向研究根据对物质的接触来计算疾病流行程度，因而得出的结果可能是患病存活率而不是患病风险 ( Szklo and

Nieto, 2000 )。另外，由于使用的调查方法不准确以及对食品摄入量变化的控制能力有限，此类研究可能在食品摄入量方面出现明显的测量误差。因此，横向研究被认为是一种“研究食品与疾病关联的相对较弱的方法” ( Sempos et al., 1999 )。出于这些原因，横向研究的结果“可能会误导，因为解释性错误很常见” ( Kraemer et al., 1005 at 103 )。横向研究的可靠性被认为低于群组追踪研究和病例对照研究 ( Greer et al., 2000 )。

*生态研究 ( ecological studies )* 比较不同群体之间的发病率。*病例报告 ( case reports )* 描述对单一被试或一小批被试的观察结果。生态研究和病例报告的可靠性在各类观察研究中最低。

## 研究综合

讨论多项研究的报告 ( 例如综述文章 ) <sup>(17)</sup> 并不就具体研究提供足够的信息，因此FDA无法据其确定关键要素，例如研究群体特征以及使用的产品的成分。同样，综述文章并不就其总结的各项研究提供详细信息，因此FDA无法确定在诸如设计、研究执行和数据分析等关键要素上这些研究是否存在缺陷。FDA必须能够审查一项研究的关键要素才能确定是否能从中得出科学结论。因此，FDA打算利用综述文章和类似出版物<sup>(18)</sup>进一步寻找可能对健康声明审查有用的研究报告，并把这些综述报告用作关于物质与疾病关系的背景信息。如果发现了更多的研究，FDA会对它们逐一评估。由于缺乏所总结的各项研究的详细信息，大多数综合分析报告 ( meta-analyses ) <sup>(19)</sup> 只会被用于进一步寻找对健康声明审查可能有用的研究报告，以及用作物质与疾病关系的背景信息。但是，对于综合分析了某特定物质与疾病关系的所有公开研究的报告，FDA会作为其健康声明审查程序的一项内容加以考虑。所综述的研究都应当与本指南列出的关键要素、质量以及其他因素一致，并应当完成适当的统计分析。

## 动物和体外研究

对于物质与疾病之间关系可能涉及的机制，FDA会把动物和体外研究作为背景信息。动物的生理机能与人类的不同。体外研究在人为的环境中进行，无法反映诸多的正常生理过程，例如影响人对食品和食品物质反应的消化、吸收、分配以及代谢

( IOM, 2005 ) 等过程。动物研究和体外研究可用于生成假设，调查假设的生物合理性，或者通过对动物食品摄入的控制来探索特定食品成份的作用机制。但是，这些研究所提供的信息不能就物质与人类疾病之间的关系作出科学结论。

## C、确定疾病风险的替代终点指标

替代终点指标是风险生物标记<sup>(20)</sup>。有研究证明这些标记能可靠地预测疾病风险，因此可在临床试验中用于测量发病率 ( Spilker, 1991 )。由于许多疾病的形成要经过很长的时期，而生物标记研究的期限较短，不一定能够发现治疗组和对照组被试之间具有有统计意义的发病率差异。

以下疾病风险替代终点指标的例子被国家卫生研究院及/或 FDA 药物评估和研究中心所接受：( 1 ) 心血管疾病的血清低密度脂蛋白 ( LDL ) 胆固醇浓度、总血清胆固醇浓度和血压；( 2 ) 骨质疏松症的骨矿物密度；( 3 ) 结肠癌的腺瘤结肠息肉；以及 ( 4 ) 二型糖尿病患者的血糖浓度升高和胰岛素抵抗力增强。

一种疾病可能有多种患病渠道，例如心血管疾病。因此，被认可的涉及某一患病渠道的替代终点指标不一定适用于涉及另一个患病渠道的某些物质。例如，长链欧米伽-3 脂肪酸通常对血清 LDL 胆固醇水平并无影响，而且研究表明，这些脂肪酸通过其他渠道改变心血管风险。因此，LDL 胆固醇水平不能用于评估长链欧米伽-3 脂肪酸与心血管疾病风险之间的关系。

## D、人体研究评估

根据本指南中描述的基于证据的评估方法，FDA 会评估每一项具体的人体研究，以确定是否可以从得出有关物质与疾病关系的科学结论。一项研究的某些关键成份——例如设计、数据收集和数据分析——可能存在严重缺陷，以致无法从研究中得出科学结论。FDA 不会使用从中无法就物质与疾病关系得出科学结论的研究，并计划把此类研究从进一步评估中剔除。以下是一些问题举例，阐释了 FDA 在关于物质与疾病关系的干预或观察研究中如何确定是否可以得出科学结论。

### **干预研究**

- *研究的被试是健康人还是患有作为健康声明对象的疾病？*健康声明声称可降低未患病人群罹患作为声明对象的疾病的发病风险。对于那些使用已患有该疾病的被试的研究，只有当其结果可以科学地扩展至未患该病的人群时，FDA 才会考虑该证据。也就是说，现有的科学根据要证明：（1）对患病人口产生的缓解或治疗作用的机制与非患病人口中减轻风险作用的机制相同，以及（2）对于患病者或健康者而言，该物质对这些机制的作用相同。如果不具备此类证据，对于那些以患有作为健康声明对象的疾病的患者为被试的研究，FDA 无法评估其物质与疾病关系，也就不能从这些研究中得出科学的结论。因此，FDA 不会使用此类研究来评估物质与疾病的关系。另一方面，举例来说，FDA 在审查关于冠心病风险降低的健康声明时，它会考虑那些对含有患有不相关疾病（例如骨质疏松症）的人或有风险（例如 LDL 胆固醇水平高）罹患冠心病的人的研究。

- *作为声明对象的疾病是否作为“主要”终点测量？*干预研究在研究开始时针对该疾病进行筛选，以最大限度地降低偏差。例如，评估结肠息肉复发的干预研究对被试进行预筛选，以确保在干预研究开始时被试不存在结肠息肉。干预研究可能会把其他疾病的结果作为次要终点评估，但并不在研究开始时对这些疾病进行筛选。例如，评估结肠息肉复发的研究也可能评估前列腺癌的发病率。但是，由于前列腺癌不是这项研究的主要终点，它就不会筛选被试以确定他们在参加研究前未患有前列腺癌。这样一来，关于前列腺癌的结果可能出现偏差，因为在研究开始时治疗组和安慰剂组之间前列腺癌患者的分布不均衡。重要的病人或疾病特征在不同研究组之间分布不均衡可能会导致错误的解读（Spilker, 1991）；因此，除非一项研究将疾病作为主要终点进行评估，否则就不能对该疾病作出科学结论。

- *研究是否包括一个适当的对照组？*一个适当的对照组由不接受物质的被试组成。如果没有一个适当的对照组，就无法确定研究终点出现的变化是来自物质或来自不相干的和未加控制的无关因素（Spilker, 1991; Federal Judicial Center, 2000）。没有一个适当的对照组，就无法得出关于物质与疾病关系的科学结论。因此，FDA 不会用此类研究来评估物质与疾病的关系。

当干预研究涉及提供一种完整的食品而不是一种食品成份时，试验食品和对照食品应当十分类似，以便评估物质与疾病之间的关系。例如，如果物质是指定的某一种

脂肪酸，那么除了该脂肪酸以外，试验食品和对照食品的其他所有食品成份上都应当相似。如果已知作为声明对象的疾病的发病风险会受到其他物质含量的影响，并且这种含量在对照食品和试验食品之间存在差异，那就无法对物质与疾病的关系得出科学结论。

- *研究的设计意图是否为测量该物质在减少疾病风险中的独立作用？*当物质是一种食品成份时，如果向干预组提供的是完整的食品或包含多种营养成份的补充剂，就有可能无法准确地认定该物质的独立作用。例如，如果声明是关于叶黄素和与年龄有关的老年黄斑变性（AMD）之间的关系，那么如果干预组食用了菠菜或包含可能对AMD有抵抗作用的其他物质（例如维生素C、维生素E和锌）的多营养成份补充剂，那么就无法从研究中得出科学结论。另一个例子是，如果物质是一种被证明能够改变血液胆固醇水平的脂肪酸，但已知会显著影响胆固醇水平的其他食品成份（例如胆固醇）的水平在干预食品和对照食品中存在显著差异，那么就不可能确定该脂肪酸的独立影响。

- *相关基线数据（例如关于替代终点指标的数据）在对照组和干预组之间是否存在显著差异？*如果被测量的终点的基线值显著不同，就很难解释干预结果。例如，在一项关于低钠食品对心血管疾病风险影响的研究中，如果干预组的基线血压水平高于对照组，就无法断定观察到的效果是否来自这两组之间钠摄入量的差别。如果交叉设计研究能在随机抽样之前有适当的时间来提供一个食品“导入期”（lead-in）<sup>(21)</sup> 或一个“洗脱期”（wash-out）<sup>(22)</sup>，就会有助于减少基线值差异的可能性。

- *对于干预组和对照组的结果如何进行统计分析？*研究数据的统计分析是一个关键因素，因为它比较提供食用物质和未食用物质的被试，以确定物质是否能够减少疾病风险。另外，在对两组以上的人进行统计分析时，数据分析应当采用适合多重比较的检验方法（例如Bonferroni、Duncan）。因此，如果没有进行对照组和干预组之间的统计分析或统计分析的方法不适当，就不能对物质在减少疾病风险方面的作用得出科学结论。因此，FDA不会把此类研究用于评估物质与疾病的关系。

- *测量哪一种疾病风险生物标记？*如上所述，如果不测量发病率或相关死亡率，替代终点指标对患病风险的测量就至关重要。如果风险生物标记不是替代终点指标，

就不能对物质和患病风险之间的关系得出科学结论（请参见上文三、C部分）。FDA 不会在评估物质与疾病关系时使用此类无法得出科学结论的研究。

- *研究进行了多长时间？*使用替代终点指标的研究应当有足够长的时间，以确保终点的任何变化是对食品干预的反应。如果研究时间很短，无法对物质的影响进行评估，就不能对物质与疾病的关系得出科学结论，因此，FDA 不会把这种研究用于评估物质与疾病的关系。例如，在评估各种饱和脂肪对血清 LDL 胆固醇浓度的干预效果时，FDA 认为至少需要三个星期的时间（Kris-Etherton and Dietschy, 1997）。

- *如果干预涉及饮食指导，是否有适当的追踪措施了解这些指导是否改变了物质摄入量？*如果饮食干预涉及提出饮食指导，而不是在控制条件下提供指定食品，就应当针对物质摄入量的变化有某种评估措施（例如进行食品评估或根据饮食指导测量摄入量生物标记）。如果对饮食指导是否造成了物质摄入量的变化没有评估措施，就不能对物质与疾病的关系得出科学结论。因此，FDA 不会把缺乏评估措施的研究用于评估物质与疾病的关系。

- *研究是在何处进行的？*受研究的群体应当与美国一般人口或拟议声明中指定的群体相关，这一点很重要。因此，FDA 要对每一项研究进行评估，以确定受研究群体营养不良或特定物质摄入量不足情况在其居住地区是否属于的普遍现象，以及/或者作为声明对象的疾病在该地区的流行与致病源是否与美国类似。在某些国家存在着的某种疾病的风险因素可能与美国的疾病风险因素无关（例如某些亚洲国家中的胃癌风险因素）。美国与研究所在国在营养、食品结构和疾病风险因素等方面的差异可能意味着该研究结果不能来推测美国总人口或美国的某一群体的情况。例如，从对营养不良的被试所作的研究那里不能得出有关营养状况相对较好的美国人的科学结论。一个营养不良的人，其营养状态和新陈代谢可能会发生严重变化，因此在一位营养不良的人和一位营养良好的人之间，物质对某一替代终点指标的影响可能存在很大差异（Shils et al., 2006）。在那些物质摄入量普遍不足的国家 and 地区从事的研究不能得出科学的结论，因为对物质摄入量的反应可能源于对营养不良的纠正，而这种结果并不是健康声明的本意。

另外，在发病原因与美国显著不同的国家或地区进行的研究也无法作出结论。例如，导致胃癌的主要风险因素（盐摄入量高和幽门螺杆菌感染）在日本的普遍程度显著高于美国。因此，就物质与胃癌的关系来说，来自日本人口的研究数据就不宜得出对美国人口潜在影响的结论。

### **观察研究**

- *收集了何种信息？*只有在证据显示物质摄入量和生物样本中的物质水平（或物质的代谢物）之间存在很强的剂量反应关系时，才可把生物样本（如血、尿、组织或头发）用于建立物质摄入量。应当有证据表明物质摄入量和生物样本中的物质或代谢水平之间（例如硒摄入量和血清硒浓度之间）存在很强的相关性<sup>(23)</sup>。如果一个特定的生物样本的相关度弱，那么从这个把该生物样本当作摄入量生物标记的研究中就不能得出科学的结论。病例对照研究中的生物样本不应当用于确定物质摄入量，因为物质的新陈代谢或浓度可能因被试患病而被改变。

- *是否用科学上可以接受并经过验证的食品评估方法来评估物质摄入量？*在评估个体平常摄入的物质方面，一次 24 小时的食用量回忆或饮食记录通常被视为不适当的评估方法，虽然这种评估可能对确定一组被试的平均摄入量有用处。建立饮食历史要求与被试进行详尽的访谈。但是，饮食历史通常也不足以用来评估物质摄入量，因为被调查者必须就一般食品及食用的份量作出判断。食用频率问卷所包括的食品数目有限，如果物质的主要来源没有包括在问卷中，就不足以评估物质摄入量。食品频率问卷往往也反映不出某一食品的不同种类或不同的烹调方法。由于这些限制，验证食品频率问卷方法对于评估食品摄入量至关重要。只有经过验证才能从科学数据中得出结论，因为缺乏验证可能会导致饮食因素与疾病或疾病相关标记之间出现假性关联<sup>(24)</sup>。

- *观察研究是否评估疾病与一种食品或食品成份之间的关系？*由于观察研究是采用记录食品摄入量的方法（例如食品频率问卷、食品回忆或饮食记录）来估算一种完整食品的摄入量，因此观察研究的一个常见的弱点是很难确知被研究群体的实际物质摄入量。另外，如果物质是一种食品成份，而不是一种完整食品，那么就还要对每一种食品中存在的食品成份含量进行另外的估算。食品成份的含量可能基于多种



因素变化，例如土壤构成、食品加工/烹调方法或储存（时间长短、温度）。因此，很难基于完整食品摄入量报告确定所食用的食品成份的准确数量。

另外，包含多种食品成份（例如多种营养品补充剂）的完整食品或产品不仅包含作为声明对象的食品成份，还包含其他的食品成份。那些食品成分可能会与被研究的食品成份的代谢或疾病及其他健康状况发病机理相关联。由于完整食品和产品（例如多种营养品补充剂）包含多种食品成份，所以很难对食品成份进行分离研究（Sempos et al., 1999）。基于完整食品记录或多种食品成份摄入量记录的研究不可能准确地认定作为声明对象的食品成份对疾病风险所产生的任何可观察到的影响是由以下因素造成的：（1）该食品成份的独立作用；（2）与其他食品成份的交互作用；（3）独立或共同起作用的其他食品成份；或（4）由于增加摄入相关食品成份含量丰富的食品而造成食品中包含的其他物质摄入量减少（请参见 Sempos et al. (1999)、Willett (1990)和 Willett(1998)关于确定某一食品中特定食品成份与疾病之间关系的复杂性）。

事实上，有证据表明，在多种情况下，基于记录常规食品摄入量的观察研究可能显示某种营养素对某种疾病的益处，然而随后进行的一项干预研究却表明，包含营养素的食品添加剂对疾病并无益处，事实上却会增加疾病风险（Lichtenstein and Russell, 2005）。例如，以往的研究报告说beta胡萝卜素含量高的水果和蔬菜与降低肺癌风险有关（Peto et al., 1981）。但是，在后来的干预研究中——Alpha维生素E和Beta胡萝卜素干预研究（ATBC）以及胡萝卜与维生素A影响试验（CARET）——显示Beta胡萝卜素补充剂会分别增加吸烟者和接触石棉的工人的肺癌风险（Alpha维生素E和Beta胡萝卜素癌症预防研究组，1994；Omenn et al., 1996）。这些研究表明，作为食品添加剂提供的一种营养素在与许多其他食品成份共同食用时会产生不同的健康影响。另外，这些研究表明仅依赖流行病学的结果可能为公众健康带来风险，因为在此类研究中对一种营养素的测量方法是仅仅记录常规食品的摄入量，并且据此认定该营养素和疾病风险之间的关系。这种影响实际上还可能有害。出于上述原因，就无法从观察研究中得出有关一种食品成份和一种疾病之间关系的科学结论。但是，观察研究可用于测量一种完整食品和一种疾病之间的关联。<sup>(25)</sup>

## E、评估研究方法的质量

对于那些在初步的评估中没有被排除的研究，FDA 会对他们分别进行研究方法质量评定。研究方法质量的评级分为高、中、低三等。FDA 会基于数种因素进行质量评级，这些因素涉及研究设计、数据收集、统计分析质量、所测量的结果类型、除去是否与美国人口有关之外被研究人群的其他特征（例如选择偏差以及规定提供重要的被试信息[例如年龄、吸烟者]）。如果科学研究充分考虑到上述所有因素或大多数因素，FDA 就会在方法质量上给予高分。FDA 会根据质量因素的不足或不确定性按程度给予中等或低等质量评分。如果研究很差，得到低等评分，就无法据以得出物质与疾病关系的科学结论。在进一步的审查中会删除这些研究。

对于到这个阶段为止在科学评估程序中依然被保留的各项研究，FDA 确定了一些评估研究方法质量的因素并列入本指南中。此类因素举例如下：

### **干预研究**

- *研究是否进行随机抽样？是否具有盲性？是否提供安慰剂？*适当的随机抽样能够消除可能影响研究结果的、除了被研究物质以外的内在及/或外在因素。当终点指标可能会受到被试意识的影响时（被试了解他或她在服用某种可能有益的物质），盲性就特别重要。当结果测量的是认知表现、心理状态（例如记忆、抑郁）或行为时，盲性至关重要。在补充剂试验中使用安慰剂可防止被试了解他/她是否在服用试验物质。
- *是否提供了入选/排除标准以及关于研究人口特征的重要信息？*例如，健康或高风险被试是否获准在研究期间服用可能影响作为声明对象的疾病的药物？如果是这样，对照组和干预组之间服药的被试比例是否相当？
- *对被试减员（被试在研究结束前退出）是否评估？是否在报告研究的文章中加以说明？减员是否合理？*如果有过多被试退出研究，则必须了解退出原因以及他们的退出对干预组和安慰剂组的人数及构成有什么影响。
- *如何验证被试遵守了研究规程？*干预研究应当包括一种机制，核查被试是否遵守了研究规程。例如，一项补充剂试验应当有一种机制确定被试服用补充剂的频率。

需要重点了解：1) 被试是服用了研究提供的所有补充剂还是仅仅服用了其中一部分，以及 2) 每一组被试中服用剂量不足的人所占的比例。

- 在进行统计分析时，使用的基线数据包含最初招收的所有被试，还是仅仅包含完成了研究的被试？如果有许多被试退出，而他们的退出对干预组和安慰组的构成产生了不同的影响，则必须确定就基线数据进行的统计分析是包含最初招收的所有被试，还是仅仅包含完成研究的被试。

- 研究是测量发病率还是测量疾病风险替代终点指标？虽然疾病风险替代终点指标已经得到验证，其结果仍然不如测量实际发病率准确。这个质量问题也可推及观察研究。

- 如何确定发病时间？当发病率是被测量的终点时，必须通过医疗记录及/或病理报告来确认作为声明对象的疾病。依赖诸如死亡证书一类较不具体的记录是不够的。这个质量问题也可推及观察研究。

### **观察研究**

- 对于疾病风险干扰因素是否做了足够的调整？一种物质与疾病关系的若干方面可能引发干扰因素。因此，对有关疾病的干扰因素作出调整很重要。这样，观察到的可能是由干扰因素造成的疾病风险不会被错误地归因于被研究的物质。例如，一种疾病可能有多种与食品无关的风险因素（例如吸烟、身体质量指数和患高血压的年龄）。因此，在评估钠和血压的关系时，就要根据年龄、吸烟、身体质量指数等对风险分析作出调整。

- 在估算食品摄入量时使用了哪一类食品评估方法？在估算“一般”食品摄入量时，经过验证的食品频率问卷的可靠性高于饮食记录或 24 小时回忆法。请参见三、B 部分。

### **F、评估科学证据的完整性**

至此，根据本指南中制定的方法，FDA会对可以从中得出科学结论的研究结果作评估，并评定作为公开信息的全部证据的力度。在做这项评估时FDA计划考虑下列因

素：研究类型（例如干预、预期群组追踪、病例对照、横向研究）、事前已评定的研究方法质量级别、各类研究的数目和样本大小、科学证据与美国人口或目标群体的相关性、支持拟议的声明的研究结果是否曾被复制<sup>(26)</sup>、以及全部证据的总体一致性<sup>(27),(28)</sup>。基于科学证据的完整性，FDA确定此类证据是否符合SSA标准，或此类证据的可信度是否足以支持关于物质与疾病关系的合格健康声明。

对每一类研究，FDA就下列因素进行审查：

- 研究数量和每组被试人数
- 方法质量（高、中或低等）。
- 每一研究类型内部的研究结果（有益影响、无影响、不利影响）。要证明一项干预研究的结果有效，干预组应当与对照组在统计上有显著不同（ $P < 0.05$ ）。至于观察研究，在数值小于或大于“1”时，风险置信区间（CI）为显著。许多研究旨在分析物质和疾病线性关系（P指趋势）的统计显著性。虽然该趋势可能具有显著性（ $P < 0.05$ ），但在不同摄入量水平上（例如摄入量的三分位、四分位或五分位）<sup>(29)</sup>的被试之间的差异不一定具有显著性。在这种情况下，研究显示没有影响。评估影响的规模（例如LDL胆固醇百分比下降）可用于比较某项研究内部的影响（例如物质的两种形式的相对影响或两种食用频率的相对影响）。
- 总体而言，各项研究之间在显示有益关系方面的一致性越高，物质与疾病关系的置信水平就越高。相互矛盾的结果并不会排除一种关联（因为研究设计要素可能导致研究中缺乏负面影响），但往往会削弱关联强度的置信水平。有益影响的规模越大，就越可能存在关联。
- 与美国总人口的相**关度**。例如，

*显示有益影响的研究在何种程度上与美国总人口或特定群体（例如老人、妇女）相关？*

*研究是否仅仅包括生活方式独特的被试（例如吸烟者、素食者）？*

*研究是否表明提供有益影响的物质的摄入量显著超出美国人通常的摄入量？*

在评估证据的完整性时，FDA 要确定这些证据是足以支持有关美国全部人口的声明，还是支持仅仅适用于某一群体的声明。如果证据仅仅支持关于某一群体的声明，在声明中就必须对此加以说明。如果物质在用于降低风险时必须显著超过美国的正常摄入量，声明中也要对此加以说明。

一般而言，干预研究为声明中所说的影响提供最强有力的证据，无论关于同一关系的现有观察研究的结果如何。干预研究的设计旨在避免选择偏差以及由于偶然因素或疾病的其他干扰因素所导致的研究发现 ( Sempos et al., 1999 )。虽然干预研究和观察研究经常同时涉及物质与疾病关系的评估，但观察研究通常不能用于排除来自更可靠的干预研究的发现 ( Sempos et al., 1999 )。单项干预研究不足以排除诸多观察研究的一致发现。但是，如果数项随机控制干预研究都显示或不显示物质与疾病的关系，则此结果会胜过任何数量的观察研究所得出的结果 ( Barton, 2005 )。这是因为干预研究的设计和控制着眼于测试物质与减少疾病风险之间的因果关系的证据，而观察研究只能识别可能的相关性。有大量的例证——例如维生素 E 和心血管疾病 ( CVD ) 之间的关系以及 beta 胡萝卜素和肺癌之间的关系——观察研究中发现的相关性被公布，但是，通过使用随机对照干预研究测试可能的相关性，没有证据表明存在这种关系 ( Lichtenstein and Russell, 2005 )。

## **G、评估明确的科学共识 ( SSA )**

明确的科学共识是指某一领域中有资质的专家对某种观点的一致程度。不同水平的科学证据连续体中，有一段是从十分有限的证据延伸到不能得出结论的证据，在这一区间上 SSA 接近完全一致。FDA 对 SSA 的认定代表 FDA 的最佳判断，即有资质的专家是否有可能一致认为，科学证据支持作为拟议健康声明对象的物质与疾病关系。FDA 将 SSA 标准作为一个强有力的标准，它为物质与疾病关系的有效性提供高度的信心。SSA 意味着该关系的有效性不太可能被新兴科学和现有科学的演进所推翻，尽管关系的确切性质可能需要进一步推敲。SSA 不要求完全一致和没有争议的科学意见。SSA 出现于可根据数据和信息作出推论的新兴科学生成之后很长的一段时间，但它的出现又处于该科学领域内就推论的有效性达成完全一致意见之前。

要使有资质的专家对一项声明的有效性达成明智的意见，与声明有关的数据和信息必须是有关科学领域中能够得到的数据和信息。因此，必须有结论说，有足够数量的、公开发表的相关科学证据显示，在不同研究和不同研究人员之间存在着一致性，据此 SSA 才能产生。可供有资质的专家使用有关证据的常用机制是由同行评审的科学杂志中出版的数据和信息。如果一位专家无法接触到所有证据，那么其意见的价值就是有限的。

在确定是否存在明确的科学共识时，FDA 也考虑到其机构外部的有资质的专家的意见，前提是此类专家做过评估并且其评估结果公开发表过。例如，FDA 会考虑：

- 由一个具有公信力的独立机构专门为此目的组成“专家小组”并记录的专家意见；
- 一个联邦政府科学机构的意见或建议，例如国家卫生研究院（NIH）、疾病控制和预防中心（CDC）或国家科学院（NAS）；
- 一个独立的专业机构的意见，例如美国儿科医生学会（AAP）营养委员会、美国心脏学会（AHA）、美国癌症学会（ACS）或由国家卫生研究院（NIH）招集的工作组或其他小组；
- 参考公开发表的对数据和信息进行批判性总结的二级科学文献。

FDA 对于联邦科学机构得出的结论给予最大的信任，特别是当此类机构对一种物质与疾病关系的有效性证据进行评审，并认为足以向公众提出饮食建议时。当联邦科学机构得出结论支持一种物质与疾病关系的效度时，FDA 通常会认为存在着明确的科学共识。其他专家机构得出的结论也可能被用于支持关于 SSA 的认定。虽然作为个人的外部专家的意见在评估 SSA 时也予以考虑，但是这种意见不足以支持达到标准的结论，特别是当联邦科学机构、专家小组或独立专业机构对同样证据进行审查后没有对该专家的意见表示支持时。外部专家或专家小组的意见最能发挥作用的情况是，有适当理由认为他们代表同一领域中更多有资质的专家的意见。最重要的是，外部专家意见的相关性取决于所检验的证据是否适用于该声明，例如对物质和疾病的描述与测量。

在得不到有资质的专家作出的结论时（例如如果支持拟议健康声明的数据相对较新，并且尚未经过一个独立专家小组或机构的审查），如果有一批令人信服的相关证据存在，FDA 也可能会认定存在着明确的科学共识。由于每个声明所涉及物质与疾病关系的性质不同，就有必要根据个案情况来考虑共识的程度以及不同意见的性质。如果用主观的定量或僵化的标准来评估科学共识，这种刻板的做法就可能使某些有效的声明遭到拒绝，尽管现有的不同意见尚缺乏说服力。

使用明确的科学共识的标准意在依靠一批健全的、相关的科学数据，以力求客观；承认支持不同的物质与疾病关系的有效性所需要的数据，其数量和类别有所差异，以力求灵活；随着时间的延伸以及研究问题和试验方法的重新界定而重新评估数据，以力求反应敏感。

## H、合格健康声明的语言规定

当一种物质与疾病关系的证据可信但未达到 SSA 标准时，就这种关系拟议的声明应当包括限定性语言，说明支持该关系的科学根据的局限程度。

健康声明语言应当具体而准确地反映科学证据的程度。但是，科学根据中的差距有时可能会限制声明中可包括的信息。例如，当科学根据有限但可信时，合格健康声明可能无法指明降低疾病风险所需要的物质摄入量。

根据 FDA 的健康声明规章，如果规章没有界定何为食品中物质的“高”含量，健康声明必须具体说明为取得所声称的影响每天必须摄入的物质数量（《美国联邦法规》第 21 卷第 101.14(d)(2)(vii)条）。在营养成分声明规章中 FDA 把“高”含量界定为食品包含物质每天摄入量的 20% 或更多（《美国联邦法规》第 21 卷第 101.54(b)条）。因此，如果尚未建立物质的每日摄入量，FDA 就无法确定何为物质的“高”含量。如果与声明相关的物质没有每日摄入量数值，FDA 会确定为达到所声明的影响每日必须的摄入量，前提是有足够的证明可以作出这种决定。请参见《美国联邦法规》第 21 卷第 101.83(c)(2)(G)条（健康声明规章：植物甾醇/甾烷醇酯和降低冠心病风险）。但是，有时关于降低风险的可信数据不够具体，FDA 不足以凭借该数据确定普通美国人所需要的摄入量。请参见 FDA 于 2004 年 9 月 8 日发表的关于 omega-3 脂肪酸和降低冠心病风险的合格健康声明的执法裁决书（Martek 申请）。

如果有可信的证据表明物质和疾病之间的关系，就有必要确定物质是在该关系中独立发挥作用，还是其作用系基于包含或取代其他物质。有时，对物质的独立作用的评估可能很困难，例如物质属于一种常规食品或大量营养素（例如脂肪或碳水化合物）。在评估一种常规食品或大量营养素可能产生的健康影响的研究中，把二者中任意一种包含在食品中通常要求排除其他常规食品或大量营养素（即取代产生相等热量的食品）。如果确认物质不是独立发挥作用及/或需要减少或包含另一种物质才能显示其益处，声明的语言就需要反映出这种情况。

## I、对现有SSA或合格健康声明的重新评估

无论是否有人提出申请，FDA都有可能重新评估一项健康声明。在重新评估时FDA会使用上文所述的科学评估程序。为了通过健康声明审查为公众健康带来最大利益，FDA会对新发布的信息进行评估，以确定基于新信息是否有必要对现有SSA或合格健康声明作出改动。例如，新的科学根据可能：（1）支持对SSA或合格健康声明的语言作出修订，（2）支持把SSA声明改为合格健康声明（QHC），或支持把QHC改为SSA声明，或（3）对作为健康声明对象的物质提出安全关注，或不再支持健康声明（SSA或QHC）。

## 四、参考资料

- 美国癌症学会，《癌症实情与图表》，2004。
- Alpha 维生素 E、Beta 胡萝卜素癌症预防研究小组。“维生素 E、Beta 胡萝卜素对男性吸烟者肺癌和其他癌症发病率的影响”，《新英格兰医学杂志》1994；330:1029-1035。
- Barton S.，“哪些临床研究提供最佳证据？最佳随机对照试验（RCT）仍然优于最佳观察研究”，《英国医学杂志》2000；321:255-256。
- Cade J、Thompson R、Burley V、Warm D，“食品摄入频率问卷的编制、验证和使用——综述”，《公共健康营养》2002；5:567-587。
- 联邦司法中心，《科学根据参考手册》第二版，2000。
- Flegal KM，“就食品和营养摄入的影响评估流行病学证据”，《美国临床营养杂志》1999；69:1339S-1344S。



- Greer N、Mosser G、Logan G、Halaas GW，“关于证据评级的一种实用方法”，《质量改进联合委员会杂志》，2000; 26:700-712。
- Hill AB，“环境和疾病：关联或因果？”《皇家医学会会议记录》，1965;58:295-300。
- Hord NG、Fenton JI，“环境最重要：探索常态和癌症前期微环境以破解食品和癌症风险的难题”，《分子营养和食品研究》，2007; 51:100-106。
- IOM，医学研究会，“食品添加剂：一种安全评估框架”，国家科学院出版社，华盛顿哥伦比亚特区，2005。
- Kraemer HC、Lowe KK、Kupfer DJ，“祝您健康：如何理解研究得出的风险结论？”牛津大学出版社，2005。
- Kris-Etherton PM、Dietschy J，“为检查个人脂肪酸对心血管疾病风险因素的影响的研究设计标准：人体和动物研究”，《美国临床营养杂志》1997; 65:1590S-1596S。
- Lichtenstein AH、Russell RM，“至关重要的营养素：食品或补充剂？”《美国医学学会杂志》2005; 294:351-358。
- Milner JA，“食品与癌症：事实与争议”，《营养与癌症》2006; 56: 216-224。
- 国家癌症研究所，《癌症术语词典》，<http://www.cancer.gov/dictionary>
- 国家研究理事会，《食品与健康：减少慢性病风险的作用》，国家科学院出版社，华盛顿哥伦比亚特区，1989。
- Omenn GS、Goodman GE、Thornquist MD、Balmes J、Cullen MR、Glass A、Keogh JP、Meyskens FL、Valanis B、Williams JH、Barnhart S、Hammer S，“beta 胡萝卜素和维生素 A 对肺癌和心血管疾病的综合影响”，《新英格兰医学杂志》1996; 334:1150-1155。
- Peto R、Doll R、Buckley JD、Sporn MB，“食品 beta 胡萝卜素能否显著减少人体癌症发病率？”，《自然》1981; 290:201-208。
- Sempos CT、Liu K、Earnst ND，“食品与营养摄入：在评估流行病学数据时应当考虑什么？”，《美国临床营养杂志》1999; 69:1330S-1338S。
- Torun B，“蛋白质能量营养失调”，摘自《健康与疾病的现代营养》，Williams 和 Williams 出版社，纽约 200。

- Spilker B , 《临床研究指南》 , 乌鸦出版社 , 纽约 , 1991.
- Subar AF、Thompson FE、Kipnis V、Midthune D、Hurwitz P、McNutt S、McIntosh A、Rosenfeld S , “对布劳克 ( Block )、惠莱特 ( Willett ) 和国家癌症研究所食品摄入频率问卷的比较验证” , 《美国流行病学杂志》 2001; 154: 1089-1099。
- Szklo M、Nieto FJ , “流行病学 : 超越基本因素” , 阿斯彭出版社 , 2000。
- Willett W.C , “营养流行病学概述” , 《营养流行病学》 , 牛津大学出版社 , 牛津 , 1990。
- Willett W.C , “分析和报告营养数据中的问题” , 摘自《营养流行病学》 , 第二版 , 牛津大学出版 , 牛津 , 1998。
- Wilson E.B , 《科学研究简介》 , 通用出版公司 , 多伦多 , 1990。

---

## 注释 :

<sup>①</sup>本指南的编写单位是美国食品与药物管理局食品安全与应用营养中心下属的营养、标签和食品添加剂办公室。

<sup>②</sup>为叙述方便,“疾病”在本文中代表“疾病或与健康有关的状况”,其定义是对人体器官、部位、结构或系统造成损坏,使之无法发挥正常作用(例如心血管疾病),或导致此类功能失常的一种健康状况(例如高血压)。《美国联邦法规》第21卷第101.14(a)(5)条。

<sup>③</sup>本新指南文件用于取代FDA发布的另一份指南,即《行业与FDA指南:过渡性基于证据的科学数据评级系统》,该指南的内容涉及对合格健康声明的科学审查。虽然该指南包括关于评估科学证据力度的章节,但本新指南文件并不包括这样的章节,因为目前正在研究消费者如何理解可用于描述健康声明证据力度的各种可能的评级系统。FDA打算重新审查其评级系统,并且在这些研究完成后发布适当指南。另外,本指南文件还取代FDA的另一份指南,即《行业指南:常规食品和食品添加剂健康声明审查中的明确的科学共识》。

<sup>(4)</sup>1997年，国会制定了《食品与药物管理现代化法》，该法建立了另一个有关健康声明的授权程序，该程序适用于基于某些联邦科学机构或国家科学院作出的权威性声明而提出的健康声明。本指南文件并不涉及该程序。

<sup>(5)</sup>请参见《行业与FDA指南：常规人类食品和人类食品添加剂标签中的合格健康声明的过渡性程序》（附件A），2003年7月10日  
(<http://www.cfsan.fda.gov/~dms/hclmgui3.html>)

<sup>(6)</sup>请参见《美国法典》第21卷第321(ff)(1)条。

<sup>(7)</sup>在《美国联邦法规》第21卷第101.14(a)(3)条中，“营养价值”指通过促进生长、补充重要营养素或提供能量等过程维持人类生命的价值。

<sup>(8)</sup>干扰因素是指与有关疾病和干预相关的因素，如果不进行控制，这些因素会妨碍研究人员得出一种干预能够产生结果的结论。

<sup>(9)</sup>自由生活的群体指根据自己的意愿选择食品和生活方式（例如吸烟、饮酒和健身）的群体。

<sup>(10)</sup>摄入量生物标记是对物质本身或生物样本中的物质代谢物（例如血清硒）的测量。经验证，这些测量能够证实该物质的摄入量。

<sup>(11)</sup>偏差属于系统性错误，可能来自被试选择的缺陷（选择偏差），也可能来自干预和疾病结果测量的缺陷（信息偏差）（Szklo和Nieto, 2000）。

<sup>(12)</sup>相对风险指有干预的人和无干预的人之间的风险比例（即发病率）。其计算方法是在前瞻群组中测量患病和未患病的被试的对物质的干预程度。经调整的相对风险可控制潜在干扰因素。

<sup>(13)</sup>病例对照研究设计的一个例子是评估与人群发病率和分布状况有关的参数，例如主要死亡原因。

(14) 发病机率指有干预和没有干预的人的发病率对比。在病例对照研究中，其计算方法是基于对物质的干预来测量被试的疾病发生。经调整的发病机率可控制潜在干扰因素。

(15) 少数横向研究是时间系列研究，用来比较不同时间段的结果（例如，在一个五年阶段内某结果的发生率在之后的另一个五年时间内是否有所变化）。

(16) 相关系数( $r$ )测量两个变量之间的相互依赖性，例如一种物质的摄入量和一种疾病的发病率。其表达方式为 -1 至+1，-1 表示完美的负相关，0 表示不相关，+1 表示完美的正相关。请参见《Webster's II New Riverside University Dictionary (Riverside Publishing Co., 1984)》。因此，相关系数越接近两端，两个变量之间的关系就越密切。

(17) 综述文章总结就某一题目进行的一些具体研究的结果。

(18) 其他例子包括书本、章节、摘要、信函和委员会报告。

(19) 综合分析系统地综合和评估已经完成或终止的临床试验的结果 (Spilker, 1991)。

(20) 风险生物标记是显示与疾病风险相关的生理状态变化的生物标记。

(21) 在随机抽样前向所有研究组提供同样食品。

(22) 交叉研究中的时间阶段，在该阶段内被试不接受干预。

(23) 相关性通过相关系数( $r$ )评估。相关系数从 -1 (负相关) 到+1 (正相关)。越接近 1，相关度越高；越接近 0，相关度越低。

(24) “食品摄入频率问卷方法的验证至关重要，因为不正确的信息可能导致饮食因素与疾病或疾病相关标记之间的错误关联。”Cade, J、Thompson, R、Burley, V、和 Warm D，“食品摄入频率问卷的编制、验证和使用——总结评估”，《公共健康营养》，5：573 页，2002 年。另请参见Subar, A., et al.,Block, Willett和“国家癌症学会食品频率问卷”，《美国流行病学杂志》，154: 1089-1099, 2001。

<sup>(25)</sup> 在 *Pearson v. Shalala* 一案中，哥伦比亚特区巡回法院指出，FDA 已经“从逻辑上确定”，没有科学根据表明服用食品添加剂中的抗氧化维生素能够降低癌症风险。现有的研究仅考察了包含抗氧化维生素的食品，而这些食品对减轻癌症风险的影响可能源于它们所包含的其他物质。法院案例 164 F.3d 650, 568 ( D.C. Cir. 1999)。但是，哥伦比亚特区巡回法院的结论是，虽然 FDA 对授予抗氧化维生素合格健康声明存有顾虑，但该问题可以通过同时发布一条醒目的免责声明来解决。免责声明可指出关于该声明的证据并不是定论，因为支持该声明的研究所基于的食品包含其他物质，而对降低癌症风险产生作用的可能实际上是这些其他物质。(同上)。法院指出，FDA 并没有声称有关食品添加剂会“威胁消费者的健康与安全。”(同上)，656 页。但是，要允许基于对个别营养素的研究提出合格健康声明还有一个更根本的问题，这个问题不象哥伦比亚巡回法院认为的那样，通过提供一条免责声明就可以解决。即使可以有把握的确定食品中某个具体成份的影响，近期对营养与食品交互反应的复杂性以及食品、生物参数和疾病之间的关系等科研结果显示，在吃一种或多种食品时某物质可能产生的健康益处会在以食品添加剂的形式摄入或通过其他食品摄入时不一定会产生。请参见 Lichtenstein 和 Russell (2005)。例如，关于食品添加剂的研究不仅证实与通过食品摄入某些营养素有关的益处会在营养素作为补充剂摄入时不一定出现，而且其中一些研究实际上表明这些营养剂反而增大了罹患它们本是用来预防的疾病的风险。(同上)。因此，一个对某个或某些食品的摄入的研究所提供的信息不足以支持营养素本身能够产生有益影响的科学结论。另外，即使营养素是通过其他食品摄入，而不是以补充剂形式摄入，生理学的影响也可能不一样，因为食品组合可能影响营养素的生物利用度和生物活性。(同上)。

<sup>(26)</sup> 对科研结果的复制对于评估科学根据的力度很重要 (Wilson, E.B, 《科学研究简介》，多佛出版社，1990；46-48 页)。

<sup>(27)</sup> 在本指南中，“一致性”是指各种得出关于某物质与疾病关系的科学结论的研究之间的一致程度。

<sup>(28)</sup> 同类和不同研究设计的结果之一致性对于评估科学根据的因果关系和力度很重要 (Hill A.B, “环境与疾病：关联或因果？《皇家医学会会议记录》1965;58:295-300)；另请参见“健康研究机构与质量、评估科学根据的系统”，该文章把“一致性”

定义为“使用类似和不同研究设计所报告的结果的类似程度。”

[\[http://www.ahrq.gov/clinic/epcsums/strengthsum.htm#Contents\]](http://www.ahrq.gov/clinic/epcsums/strengthsum.htm#Contents)

<sup>(29)</sup>摄入量的三分位、四分位和五分位是把研究人口划分为三组、四组或五组，使物质的平均摄入量在不同组别间形成差异（例如摄入量最低的组代表摄入量最低的三分之一，摄入量最高的组代表最高的三分之一）。在划分研究群体时要使每一组拥有同样数目的被试。

---

\*更新的网页：

- [对 2004 年 1 月 27 日健康声明申请的答复：绿茶和降低癌症风险健康声明](#)  
2005 年 6 月 30 日
- [关于“西红柿和前列腺、卵巢、胃和胰腺癌（美国人长寿陈情书）”的信函](#)  
2005 年 11 月 8 日
- [关于钙和结肠/直肠、乳房和前列腺癌以及复发性结肠息肉的信函](#) 2005 年 10 月 12 日

---

本文件取代 2007 年发布的[草稿](#)。