

行業指南：用於健康聲明科學評估的基於證據的審查系統（終稿）

2009 年 1 月

包含非約束性建議

欲索取本文件，請與以下部門聯絡：

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

(電話) 301-436-1450

<http://www.cfsan.fda.gov/guidance.html>

您可以在任何時候以書面或電郵形式對本指南提出意見。請把有關指南的書面意見發送至：Division of Dockets Management (HFA-305), Food and Drug Administration, 5630 Fishers Lane, rm. 1061, Rockville, MD 20852。電子形式的意見可發送至：<http://www.regulations.gov>。所有意見都應當註明文件號碼，該文件號碼載於「聯邦公告」（*Federal Register*）中的可用性通知。

美國衛生與公眾服務部

食品與藥物管理局

食品安全與應用營養中心

2009 年 1 月

包含非約束性建議

目錄

1. [引言](#)
2. [背景資訊](#)
3. [用於健康聲明科學評估的基於證據的審查系統](#)
4. [參考資料](#)

包含非約束性建議

行業指南^[1]

用於健康聲明科學評估的基於證據的審查系統

本指南代表食品與藥物管理局 (FDA) 對該题目的當前意見，並不建立或賦予任何人以任何權利，也並不對 FDA 或公眾具有約束力。您可以採用替代方法，但該方法必須滿足有關法律和法規的要求。您如果希望討論一種替代方法，請與負責實施本指南的 FDA 人員聯絡。您如果無法確定應該與 FDA 的哪些人員聯絡，請撥打本指南封面上列出的有關電話號碼。

一、引言

本文件為企業提供指南。它代表 FDA 目前對以下問題的意見：1) 評估健康聲明科學證據的程序，2) 「聯邦食品、藥物與化妝品法」(簡稱「法案」) 403(r)(3) 款 (21 U.S.C. 343(r)(3)) 和 21 CFR 101.14(c) 中顯著的科學一致性 (SSA) 標準的含義，以及 3) 用於支持一項合格健康聲明 (QHC) 的可信科學證據。

本指南文件描述了一種基於證據的審查系統，FDA 打算把該系統用於涉及一種物質與某種疾病或健康問題之關係的合格健康聲明或 SSA 健康聲明，評估其公開的科學

證據。⁽²⁾本指南文件說明了FDA目前對其應當使用的科學審查方法的意見，旨在為健康聲明申請人提供指南。⁽³⁾

本指南文件中陳述的具體題目是：(1) 確定評估物質/疾病關係的研究，(2) 確定疾病風險替代終點 (surrogate endpoints)，(3) 評估人體研究，以確定是否能從中得出有關物質/疾病關係的結論，(4) 評估可從中得出物質/疾病關係科學結論的每一項人體研究的方法品質，(5) 評估科學證據的完備性，(6) 評估顯著的科學一致性，(7) 合格健康聲明的語言針對性，以及(8) 重新評估現有 SSA 或合格健康聲明。

FDA 的指南文件 (包括本指南) 並不規定在法律上必須履行的責任。相反，指南文件只是描述 FDA 對某一題目的當前意見，只應當視作建議，除非引用具體監管或法律要求。在 FDA 指南文件中，「應當」一詞是指提出建議，而不是強制性規定。

二、背景資訊

「1990 年營養標籤和教育法」 (NLEA) (Pub. L. 101-553) 旨在促進為消費者提供關於食品的更具有科學性的資訊。除其他條款外，NLEA 指示 FDA 發佈規章，要求在食品 (包括食品補充劑) 標籤中包含描述物質/疾病關係的聲明 (「健康聲明」)，但此等聲明須經 FDA 事先審查和授權。⁽⁴⁾就這些健康聲明——即關於物質/疾病關係的聲明——而言，FDA 透過規章把「物質」定義為一種具體的食品或食品成份 (21 CFR 101.14(a)(2))。經授權的健康聲明既可用於常規食品又可用於食品補充劑，前提是產品中的物質和產品本身符合授權規章中的有關標準。健康聲明針對公眾或特定群體 (例如老年人)，旨在幫助消費者保持健康的飲食習慣。

在評估一項經授權健康聲明的申請時，FDA 考量作為聲明主題的關係之證據是否符合 SSA 標準。該標準源於 21 U.S.C. 343 (r)(3)(B)(i)，它規定 FDA 應當授權一份用於常規食品的健康聲明，前提是 FDA 「基於所有公開資訊 (包括設計健全、符合公認的科學程序和原則的研究提供的證據) 認定，在受過科學訓練、擁有科學

經驗、有資格評估此類聲明的專家中存在相當普遍的共識，認為該聲明受到此類證據的支持。」法律規定，常規食品健康聲明一律採用該科學標準；FDA 則以規章形式決定對食品補充劑健康聲明也採用同樣標準。請參見 21 CFR 101.14(c)。

合格健康聲明起源於上訴法院對 *Pearson v. Shalala (Pearson)* 一案的裁決。在該案件中，原告對 FDA 不授權把健康聲明用於食品補充劑標籤中的四種具體的物質疾病關係提出異議。雖然地區法院裁定 FDA 勝訴 (14 F. Supp. 2d 10 (D.D.C. 1998))，但哥倫比亞特區美國巡迴上訴法院推翻了下級法院的裁決 (164 F.3d 650 (D.C. Cir.1999))。上訴法院認為，第一修正案並不許可 FDA 拒絕它認為可能誤導公眾的健康聲明，除非 FDA 基於合理證據認定透過免責聲明無法消除可能存在的欺騙。上訴法院還認為，「行政訴訟法」(APA) 要求 FDA 澄清授權健康聲明時採用的「顯著的科學一致性」(SSA) 標準。

1999 年 12 月 22 日，FDA 宣佈發佈「行業指南：常規食品和食品補充劑健康聲明審查中的顯著的科學一致性」(64 Fed. Reg.17494)。該指南文件的發佈旨在根據上訴法院在 *Pearson* 一案中的第二項意見澄清 FDA 對 SSA 標準的解釋。

2002 年 12 月 20 日，FDA 宣佈它打算擴大執行 *Pearson* 案裁決的範圍，把常規食品健康聲明也包括在內 (67 Fed. Reg. 78002)。FDA 認識到有必要建立一個合格健康聲明的科學框架，因此成立了「改善消費者營養健康資訊」(Consumer Health Information for Better Nutrition) 工作組。該工作組認識到，如果公眾能夠獲得和使用常規食品以及食品補充劑標籤中的更多和更好的資訊，更明智地作出購買食品的決定，就可能獲得重大的健康益處，因為這些資訊不僅包含價格、便利程度和口味，而且包含基於科學的健康說明。消費者如果掌握更多關於食品和食品補充劑對於健康的益處的科學資訊，就能夠降低諸多慢性病風險，對其長期健康產生有益影響。

為了最大限度地發揮 FDA 聲明審查程序對增進公眾健康的作用，工作組在「最終報告」中^⑤提出了一種程序，用於根據個案對所有完整的申請進行排序，排序時基於數種因素，包括作為申請主題的食品或食品補充劑是否可能對一種嚴重的或危及生命的疾病產生重大影響；證據的可信度；消費者調查是否顯示聲明不具有誤導性；作為聲明主題的物質是否經過 FDA 安全審查 (即是否屬於經授權食品添加劑，是否

屬於「公認安全」(GRAS) 物質，是否列入清單，或是否就GRAS通知收到「不反對」信函)；作為聲明主題的物質是否予以適當定性，以便評估現有研究的相關性；疾病是否根據一個公認的合格專家機構制定的公認標準加以定義和評估；一個公認的合格專家機構是否曾經對證據或聲明進行過審查。

作為工作組最終報告的部分內容，FDA 建立了一個基於證據的過渡性審查系統，並打算把該系統用於評估作為合格健康聲明主題的物質/疾病關係。在審議 1999 年 12 月 22 日 SSA 指南文件和 2003 年工作組報告時，FDA 明確認識到 SSA 健康聲明與合格健康聲明的科學審查程序之要素十分相似。由於這兩類聲明的科學審查程序相似，FDA 打算把本指南中闡述的方法，用於評估就 SSA 健康聲明或合格健康聲明提出的申請中包含的科學證據。本指南中闡述的基於證據的審查系統將有助於 FDA 確定科學證據是否符合 SSA 標準，如果不符合，則確定證據是否支持一項合格健康聲明。除一項科學審查外，健康聲明還需要經過監管審查。符合 SSA 標準的健康聲明的授權方式是在「聯邦公告」中發佈一項最終決定或過渡性最終決定。至於有可信證據支持的合格健康聲明，FDA 會發出一封信函，表明其有意考慮執行裁量權 (enforcement discretion)。

針對上訴法院在 *Pearson* 一案中的第二項裁決，FDA 頒佈了「行業指南：常規食品 and 食品補充劑健康聲明審查中的顯著的科學一致性」(64 Fed. Reg. 17494)，以澄清 FDA 對 SSA 標準的解釋。雖然本指南取代了該文件，但 FDA 認為本指南依然符合法院的裁決，1999 年指南中闡述的 SSA 基本原則沒有改變，關於 SSA 的決定仍然要求 FDA 作出最佳判斷，確定合格專家是否有可能認同作為擬議健康聲明主題的物質/疾病關係的科學證據。事實上，本指南中對 SSA 的許多說明與 1999 年指南中的語言完全相同。本指南反映了 FDA 在審查科學證據的工作中對進一步科研成果的利用，而不是改變它對 SSA 內容的理解。

三、用於健康聲明科學評估的基於證據的審查系統

A、何為基於證據的審查系統？

基於證據的審查系統指對於支持一項聲明的證據的可信度進行系統性科學評估。就健康聲明而言，該系統評估支持一項關於物質/疾病關係的擬議聲明的科學證據的

可信度。該評估程序涉及一系列步驟，包括確定需要評估的科學研究和其他資料，排除無法從中得出物質/疾病關係結論的資料，對剩餘的研究進行方法品質評估，並且評估科學證據的完整性，具體做法是審查研究類型、方法品質、支持和反對聲明的證據的數量（考慮到各類研究的數目和研究樣本規模）、與美國人口或目標群體的相關性、支持擬議聲明的研究結果的可複製性、以及證據的總體連貫性。在審查科學證據的完整性之後，FDA 確定是否存在支持一項經授權健康聲明的顯著的科學一致性，或是否存在支持一項合格健康聲明的可信證據。

B、確定評估物質/疾病關係的研究

FDA 審查涉及物質/疾病關係的公開資料和書面資訊，包括健康聲明申請中必須提及的各項研究（21 CFR 101.70）。透過文獻搜索，FDA 確定與擬議健康聲明相關的其他研究。在審查物質/疾病關係證據的可信度之前，FDA 把關於人體研究的個別相關文章與其他類型的資料和資訊區別開來。FDA 打算把審查的重點放在報告人體干預和觀察研究的文章，因為只有此類研究能夠提供就物質/人類疾病關係得出科學結論的證據。其次，FDA 將在審查科學證據過程中考量若干關鍵性問題：

- *研究是否指明和測量作為聲明主題的物質？* 研究應當指明一種可以測量的物質。「物質」的定義是一種具體食品或食品成份，無論是常規食品還是食品補充劑。21 CFR 101.14(a) (2)。例如，食品成份可以是一種營養品或食品配料。⁽⁶⁾如果物質作為常規食品成份按照遞減營養級別（decreased dietary levels）消費，則物質必須是一種營養品，而且這種營養品必須包括在「營養事實」標籤中（21 CFR 101.14(b)(2)）。如果物質不是按照遞減營養級別消費，則物質必須有助於口感、香味、營養價值，⁽⁷⁾或產生 21 CFR 170.3(o)中列舉的某種對食品的技术影響，並且必須安全，足以按照合法方式證明聲明的合理性（21 CFR 101.14(b)(3)）。
- *研究是否適當指明和測量作為聲明主題的具體疾病或健康問題？* 「疾病或健康問題」的定義是對人體的器官、組成部分、結構或系統造成損害，使之不能發揮正常功能（例如心血管疾病）或導致此類功能失常的健康狀態（例如高血壓）。21 CFR 101.14(a) (5)。研究應該藉由測量發病率、相關死亡率或經驗證可預測具體疾病風險的替代終點，指明一種具體可測量的疾病或健康問題。

例如，癌症包含 100 多種不同的疾病，每一種疾病都表現為異常細胞的過度成長和擴散（美國癌症學會，2004）。癌症基於受影響的器官和組織部位分為不同的疾病類別（國家癌症研究所）。處於不同器官部位的癌症有不同的風險因素、治療方式和死亡風險（美國癌症學會，2004）。遺傳和環境（包括食品）因素都會影響不同類別的癌症的風險，風險因素可包括每一類癌症的家族病史、吸煙、飲酒、超重或肥胖、接觸紫外線或電離子輻射、接觸致癌化學品、食品因素等。每一類癌症的病因、風險因素、診斷和治療方法各不相同（Hord et al., 2007; Milner et al., 2006）。由於每一類癌症基於器官位置、風險因素、治療選擇和死亡風險都是一種獨特的疾病，FDA當前的方針是評估健康聲明或合格健康聲明申請中的每一種癌症，以確定就該類癌症是否存在支持物質/疾病關係的科學證據，這種癌症構成 21 CFR 101.14(a)(5)所規定的疾病。FDA遵循此一方針發佈了數封執行裁量權信函，包括[2005 年 6 月 30 日發佈的「綠茶和癌症」](#) [*](#)、[2005 年 11 月 8 日發佈的「西紅柿/番茄紅素和各類癌症」](#) [*](#)、[2005 年 11 月 12 日發佈的「鈣和各類癌症」](#) [*](#)、以及一份載於「聯邦公告」的通知，標題是「健康聲明與合格健康聲明；食品脂肪和癌症、大豆蛋白質與冠心病、抗氧化維他命與某些癌症以及硒和某些癌症；重新評估」（72 Fed. Reg. 72738, 2007 年 12 月 21 日）。

在對這些關鍵性問題進行審議後，FDA 把各項研究進行了分類。

干預研究

在干預研究中，為受試者提供相關（干預組）物質（食品或食品成份），通常是採取常規食品或食品補充劑形式。應當對物質的品質和數量進行控制。在隨機對照試驗中，受試者進行隨機編組。個體受試者之間的情形可能並不相似，但在隨機抽樣後，干預組和對照組的情形應當類似。隨機對照試驗提供對物質和疾病之間的因果關係的最佳評估，因為已知混雜因素（即可能影響疾病風險的其他因素）受到控制。透過把受試者按照隨機方式分配至干預組和對照組，此類研究避免了選擇偏差——即最有可能產生有益結果（無論是否干預）的受試者被優先選入干預組。降低可能的偏差可以透過「盲性」研究，即受試者不知道自己是否接受干預，或是

「雙盲」研究，即受試者和評估結果的研究人員都不知道誰在干預組，誰在對照組。透過控制試驗環境——包括消費的物質數量和成份——以及所有其他飲食因素，此類研究還能夠最大限度地減少變量或混雜因素對結果的影響。⁽⁸⁾因此，隨機對照干預研究提供最強有力的證據，能夠說明物質和疾病之間是否存在關係(Greer et al., 2000)。

另外，此類研究還能夠就干預和結果之間的因果關係，提供令人信服的證據 (Kraemer et al., 2005 at 113)。但是，隨機分組可能導致對照組和治療組之間受試者特徵分佈不均勻 (例如，基線年齡或血液[血清或血漿]LDL 膽固醇水準顯著不同)。如果基線值顯著不同，那麼就很難確定研究結束時的差異是源於干預還是源於研究開始時業已存在的差別。當物質作為補充劑提供時，應當為對照組提供安慰劑。當物質是食品時，可能無法提供安慰劑，因此這種研究中的受試者可能不具有盲性。雖然在這種情形下研究不具有盲性，仍然需要有一個對照組才能從研究中得出結論。

隨機對照試驗通常採用平行設計或交叉設計。平行設計研究涉及兩組受試者，測試組和對照組，兩組同時接受物質或互為對照。在交叉設計中，在一個指定階段結束後，所有受試者在干預組和對照組之間互換。

雖然干預研究是確定因果關係的最可靠的研究類別，把基於選定群體的研究結果推廣至不同的群體可能不具有科學可靠性。例如，如果提供證據的研究表明攝取一種物質和降低青少年糖尿病風險之間相關，則不應當基於此類研究結果推斷成年人的糖尿病風險。

觀察研究

觀察研究測量物質和疾病之間的關聯。觀察研究缺乏干預研究所具有的環境控制。觀察研究最能夠反映自由生活⁽⁹⁾的群體，可能建立物質和疾病之間的關聯。與干預性研究相反，觀察研究不能確定觀察到的一種關係，是反映物質導致疾病風險降低的關係，還是一種巧合 (Sempos et al., 1999)。由於在研究開始時沒有基於各種疾病風險因素對受試者進行隨機分組，因此需要收集和調整疾病風險的已知混雜因素，以便最大限度地降低偏差。例如，應當收集關於每一位受試者風險因素的資訊

(例如年齡、種族、體重和是否抽煙)並用於調整資料，以便對物質/疾病關係作出準確測量。需要調整的風險因素針對研究的每一種疾病確定。例如，心血管疾病的風險隨著年齡增大；因此，需要對年齡作出調整，以便排除可能的混淆。

在確定作為聲明的主題的物質是否得到適當測量時，有必要重點審查食品攝入估算方法。許多觀察研究依賴受試者自己報告的飲食(例如飲食記錄、24小時回憶、飲食歷史和食用頻率問卷)，這些報告只是對攝入食品的大致估計(國家研究理事會，1989)。飲食記錄所基於的前提是食品重量能夠準確反映食品攝入量。受試者把他們食用的食品過秤，記錄其重量。24小時回憶方法要求受試者描述在過去24小時內他們食用的食品以及每種食品的份量。食品歷史方法利用問卷或訪談估計受試者在某一階段內的普通飲食。食用頻率問卷是在大型食品和健康觀察研究中最常用的食品估算工具。經驗證的食用頻率問卷在評估「通常」食品攝入量方面的可靠性，高於飲食記錄方法或24小時回憶方法(Subar et al., 2001)。該問卷讓參加者報告食用的頻率和份量，報告的食品是從一個食品清單中選擇，報告時間是一個指定的階段。上文描述的食品估算方法的問題之一是在自我報告某些食品時可能會出現偏差。例如，超重的人傾向於少報其食品份量(Flegal et al., 1999)，因此對實際食用的物質常常低估，如果某種物質有可靠的攝入生物標記⁽¹⁰⁾，則常常對這些生物標記進行測量，而不採用自我報告攝入量的方法。

觀察研究可具有前瞻性或追溯性。此類研究可能出現不同形式的偏差(資訊與選擇)。⁽¹¹⁾在前瞻性研究中，研究人員招收受試者，在疾病結果發生之前觀察他們。前瞻性觀察研究把發病率與物質食用狀況進行比較。在追溯性研究中，研究人員在疾病發生後審查受試者的病歷及/或訪談受試者。追溯性研究特別容易出現測算錯誤和回憶偏差，因為它們依賴受試者對過去攝取的食物回憶。由於觀察研究控制變量的能力有限，常常會受到混雜因素的干擾，例如複雜的物質/疾病反應。

設計得當的觀察研究能夠提供有用的資訊，發現可能的關聯，以便透過干預研究確認(Kraemer et al., 2005 at 107)。與干預研究相反，即使設計最好的觀察研究也不能建立干預和結果之間的因果關係(Kraemer et al., 2005 at 114)。但是，如上所述，干預研究則能夠測試是否有證據表明物質和降低的疾病風險之間存在因果關係。在某些情形下，從中可以得出科學結論的觀察研究可用於支持SSA或合

格健康聲明中的物質/疾病關係。如下文所述，每一種觀察研究設計都有其長處和缺點。(Sempos et al., 1999)。

隊列研究 (cohort studies) 屬於前瞻性研究，此類研究比較食用某一作為聲明主題的物質的受試者之發病率與沒有食用這種物質的受試者的發病率。由於物質的攝入先於發病，這種研究設計確保受試者食用物質不是由於生病。隊列研究能夠生成相對風險估計 (Szklo and Nieto, 2000)。(12) 隊列研究被認為是最可靠的觀察研究設計 (Greer et al., 2000)。

在*病例對照研究 (case-control studies)* 中，對有病的受試者 (病例組) 與沒有病的受試者 (對照組) 進行比較。(13) 對病例組和對照組都採用食品估算方法對以往的物質攝入量進行估算。此類追溯性研究常常問及在疾病確診前至少一年的食品攝入狀況，因此很難獲得準確的攝入量估值。另外，一項關鍵的假設是食品攝入沒有因疾病的進程或瞭解到有病而改變。因此，病例對照研究的設計並不控制因疾病引起或作為反應的攝入量變化。病例對照研究能夠生成發病概率，這是對得病相對風險的估計 (Szklo and Nieto, 2000)。(14) 病例對照研究的可靠性被認為低於隊列研究 (Greer et al., 2000)。

巢式病例對照研究 (nested-case control studies) 或*病例隊列研究 (case-cohort studies)* 使用從預先定義的隊列中挑選的受試者，例如一個正在進行的隊列研究群體。病例是隊列中被診斷患病 (例如肺癌) 的受試者。在巢式病例對照研究中，對照組是每一次一個病例 (例如肺癌) 被診斷時，從有風險的個人中選擇的受試者。在病例隊列研究中，對照組從基線隊列中隨機選擇 (Szklo and Nieto, 2000)。在此類研究中，可以計算相對風險或發病概率。巢式病例對照研究或病例隊列研究的可靠性被認為低於隊列研究，但高於病例對照研究。

橫斷面研究 (cross-sectional studies) 通常收集在某一時間點上患病者和無病者攝入食品的資訊。(15) 此類研究可用於確定可能的相關性 (即確定一種物質攝入量和發病率之間的相關係數(16))，以及為後來的前瞻性研究提供基線資訊 (Kraemer et al., 2005 at 99-100)。但是，由於食品攝入量和疾病狀況同時測量，不可能確定物質的攝入是影響疾病風險的因素還是患病的結果。橫斷面研究基於暴露風險計算發病率，因而得出的結果可能是病人存活率，而不是發病風險 (Szklo and Nieto,

2000)。另外，由於使用的調查方法不準確以及對食品攝入量變化的控制能力有限，此類研究可能在食品攝入量方面出現重大估算誤差，因此，橫斷面研究被認為是一種「研究食品/疾病關聯的相對較弱的方法」(Sempos et al., 1999)。出於這些原因，橫斷面研究的結果「可能誤導，因為解釋性錯誤很常見」(Kraemer et al., 1005 at 103)。橫斷面研究的可靠性被認為低於隊列研究和病例對照研究(Greer et al., 2000)。

生態研究 (ecological studies) 比較不同群體之間的發病率。*病例報告 (case reports)* 描述對單一受試者或一小批受試者的觀察結果。生態研究和病例報告的可靠性在各類觀察研究中最低。

調查合成研究

討論多項研究的報告(例如評論文章) (17) 並不就具體研究提供足夠的資訊，因此FDA無法據其確定關鍵要素，例如研究群體特徵以及使用的產品的成份。同樣，評論文章並不就其總結的各項研究提供詳細資訊，因此FDA無法確定在關鍵要素(設計、研究執行和資料分析) 方面這些研究是否存在缺陷。FDA必須能夠審查一項研究的關鍵要素才能確定是否能從中得出科學結論。因此，FDA打算利用評論文章和類似出版物(18) 尋找可能對健康聲明審查有用的進一步研究報告，並用作關於物質/疾病關係的背景資訊。如果發現進一步研究，FDA打算對它們逐一評估。由於缺乏所總結的各項研究的詳細資訊，大多數綜合分析(meta-analyses) (19) 只會被用於尋找對健康聲明審查可能有用的進一步研究報告，以及用作物質/疾病關係的背景資訊。但是，如果一份綜合分析評估了就特定物質/疾病關係發表的所有公開研究，FDA即會將該綜合分析作為其健康聲明審查程序的一項內容。所評估的研究應當與本指南列出的關鍵要素、品質以及其他因素一致，統計分析應當健全。

動物和體外研究

至於物質和疾病之間的任何關係可能涉及各種機制，FDA 打算把動物和體外研究用作背景資訊。動物生理學與人類生理學不同。體外研究是在一種人為的環境中進行，不能反映眾多的正常生理學程序，例如消化、吸收、分佈以及影響人對食品 and 食品物質反應的新陳代謝(IOM, 2005)。動物研究和體外研究可用於生成假

設，調查假設的生物合理性，或者透過有控制的動物食品攝入探索特定食品成份的作用機制；但是，這些研究所提供的資訊，不能用於對物質和人類疾病之間的關係得出科學結論。

C、確定疾病風險的替代終點

替代終點是風險生物標記⁽²⁰⁾，這些標記被研究證明是疾病風險的可靠預測變量，因此可在臨床試驗中用於測量發病率 (Spilker, 1991)。由於許多疾病是在長期內形成，不一定能夠透過足夠長時間的研究，以發現治療組和對照組受試者中發病率的有意義的統計差異。

以下疾病風險替代終點的例子被國家衛生研究院及/或 FDA 藥物評估和研究中心所接受：(1) 就心血管疾病而言，血清低密度脂蛋白 (LDL) 膽固醇濃度、總血清膽固醇濃度和血壓；(2) 骨質疏鬆症的骨礦物密度；(3) 結腸癌的腺瘤結腸息肉；以及 (4) 二類糖尿病患者的高血糖濃度和胰島素抗性。

一種疾病可能有多個路徑，例如心血管疾病。因此，涉及單一路徑的被接受的替代終點，不一定適用於另一個路徑所涉及的某些物質。例如，長鏈歐米伽-3 脂肪酸通常對血清 LDL 膽固醇水準並無影響，研究表明，這些脂肪酸透過一個不同的路徑改變心血管風險。因此，LDL 膽固醇水準不能用於評估長鏈歐米伽-3 脂肪酸和心血管疾病風險之間的關係。

D、人體研究評估

根據本指南中描述的基於證據的評估方法，FDA 打算評估每一項具體的人體研究，以確定是否可以從中得出有關物質/疾病關係的科學結論。一項研究的某些關鍵成份——例如設計、資料收集和資料分析——可能存在嚴重缺陷，以致無法從研究中得出科學結論。FDA 不打算使用從中無法就物質/疾病關係得出科學結論的研究，並計劃把此類研究從進一步評估中刪除。以下是一些問題舉例，闡釋了 FDA 在關於物質/疾病關係的干預或觀察研究中，如何確定是否可以得出科學結論。

干預研究

- *研究中受試者是健康人還是患有作為健康聲明主題的疾病？*健康聲明涉及降低並不患有作為聲明主題的疾病的人的發病風險。如果研究對象是患有作為聲明主題的疾病的受試者，FDA 只是在科學研究認為研究結果可以適當應用至沒有患病的人士時，才會考量其研究證據。換言之，現有的科學證據要證明：（1）在患病人口中測得的減輕或治療影響的機制，與非患病人口中減輕風險影響的機制相同，以及（2）對於患病者或健康者而言，物質對這些機制的影響相同。如果研究是使用患有作為健康聲明主題的疾病的人來評估物質/疾病關係，而上述兩種證據又不可得，FDA 就不能夠從這些研究中得出科學結論，因此，FDA 不會使用此類研究以評估物質/疾病關係。另一方面，舉例來說，如果 FDA 審查關於冠心病風險降低的健康聲明，它會考量包括患有不相關疾病（例如骨質疏鬆症）的人，或有可能（例如 LDL 膽固醇水準高）患上作為聲明主題的疾病的人的研究。

- *作為聲明主題的疾病是否作為「主要」終點測量？*干預研究在研究開始時對疾病流行病例進行篩選，以最大限度地降低偏差。例如，評估結腸息肉復發的干預研究對受試者進行預篩選，以確保在干預研究開始時不存在結腸息肉。干預研究可能把其他疾病的結果作為次要終點評估，但並不在研究開始時基於這些疾病進行篩選。例如，評估結腸息肉復發的研究也可能評估前列腺癌的發病率，但是，由於前列腺癌終點不是主要終點，這項研究不會對受試者進行篩選，以確定他們在參加研究前不患有前列腺癌。因此，關於前列腺癌的結果可能出現偏差，其原因是在研究開始時，治療組和安慰組之間前列腺癌患者的分佈不均衡。重要的病人或疾病特徵在不同研究組之間的分佈不均衡，可能導致錯誤的解釋（Spilker, 1991）；因此，除非一項研究把結果作為主要終點進行評估，否則不能就該疾病終點得出科學結論。

- *研究是否包括一個適當的對照組？*一個適當的對照組由不接受物質的受試者組成。如果沒有包含一個適當的對照組，就無法確定有關終點的變化，是來自物質還是來自不相關的和未加控制的外部因素（Spilker, 1991; Federal Judicial Center, 2000）。沒有一個適當的對照組，就無法得出關於物質/疾病關係的科學結論，因此，FDA 不打算把此類研究用於評估物質/疾病關係。

當干預研究涉及提供一種完整的食品而不是一種食品成份時，試驗食品和對照食品應當十分類似，以便評估物質和疾病之間的關係。例如，如果物質是一種特定類別

的脂肪酸，那麼試驗食品和對照食品的組成在所有食品成份上應當類似，只有該特定的脂肪酸例外。如果已知作為聲明主題的疾病的發病風險，受到其他可能影響該疾病風險的物質影響，並且這種含量在對照食品和試驗食品之間存在差異，就無法對物質/疾病關係得出科學結論。

- *研究的設計意圖是否在於測量減少疾病風險的物質的獨立作用？*當物質是一種食品成份時，如果向干預組提供完整食品或包含多種營養成份的補充劑，則可能無法準確地確定是否屬於獨立影響。例如，如果聲明是關於葉黃素和與年齡有關的老年黃斑變性（AMD）之間的關係，則如果干預組食用菠菜或包含可能對AMD有抵抗作用的其他物質（例如維他命C、維他命E和鋅）的多營養成份補充劑，那麼就無法從研究中得出科學結論。另一個例子是，如果物質是一種被證明能夠改變血液膽固醇水準的脂肪酸，但已知會顯著影響膽固醇水準的其他食品成份（例如膽固醇）的水準，在干預食品和對照食品中存在差異，那麼就不可能確定該脂肪酸的獨立影響。

- *相關基線資料（例如關於替代終點的資料）在對照組和干預組之間是否存在顯著差異？*如果被測量的終點的基線值顯著不同，則很難解釋干預結果。例如，在一項關於低鈉食品對心血管疾病風險影響的研究中，如果干預組的基線血壓水準高於對照組，就無法斷定觀察到的影響是否來自這兩組之間鈉攝入量的差別。在隨機抽樣之前，如果在足夠長的時間內為交叉設計研究提供一個食品「導入期」（lead-in）[\(21\)](#)或一個食品「洗脫期」（wash-out）[\(22\)](#)，就有助於減少基線值差異的可能性。

- *對干預組和對照組的結果如何進行統計分析？*研究資料的統計分析是一項關鍵因素，因為它提供食用物質和不食用物質的受試者之間的比較，以確定物質是否能夠減少疾病風險。另外，在對兩組以上的人進行統計分析時，資料分析應當採用適合多方比較的測試（例如Bonferroni、Duncan）。因此，如果沒有進行對照組和干預組之間的統計分析或統計分析方法不適當，就不能對物質在減少疾病風險方面的作用得出科學結論，因此，FDA不打算把此類研究用於評估物質/疾病關係。

- *測量哪一種疾病風險生物標記？*如上所述，當不測量發病率或相關死亡率時，替代終點對測量風險就至關重要。如果風險生物標記不是替代終點，就不能對物質和

疾病風險之間的關係得出科學結論（請參見上文三、C 部分）。FDA 不打算在評估物質/疾病關係時使用此類無法得出科學結論的研究。

- *研究進行了多長時間？*使用替代終點的研究應當有足夠長的時間，以確保終點的任何變化是對食品干預的反應。如果研究時間很短，無法對物質的影響進行評估，就不能對物質/疾病關係得出科學結論，因此，FDA 不打算把這種研究用於評估物質/疾病關係。例如，在評估各種飽和脂肪對血清 LDA 膽固醇濃度的干預效果時，FDA 認為至少需要三個星期的時間（Kris-Etherton and Dietschy, 1997）。

- *如果干預涉及飲食要求，是否有適當的追蹤措施瞭解這些要求是否導致物質攝入量的改變？*如果飲食干預涉及提出飲食要求，而不是在控制條件下提供指定食品，應當有某種評估措施，瞭解物質攝入量的變化（例如進行食品評估或根據飲食要求測量攝入量生物標記）。如果沒有評估措施，不瞭解飲食要求是否導致物質攝入量的變化，就不能對物質/疾病關係得出科學結論，因此，FDA 不打算把缺乏評估措施的研究用於評估物質/疾病關係。

- *研究在何處進行？*研究人口應當與美國總人口或擬議聲明中指定的群體相關，這一點很重要。因此，FDA 對每一項研究進行評估，以確定在研究人口居住的地區營養不良或特定物質攝入量不足是否屬於普遍現象，及/或作為聲明主題的疾病的普遍程度或發病原因是否與美國類似。某些國家可能存在一些關於某種疾病的風險因素，但這些風險因素與美國的疾病風險因素無關（例如某些亞洲國家中的胃癌風險因素）。美國和研究所在國在營養、食品結構和疾病風險因素等方面的差異，可能意味著研究結果不能應用至美國人口或某一群體。例如，如果受試者營養不良，則不能基於對他們的研究得出有關營養相對較好的美國人口的科學結論。當一個人營養不良時，營養狀態和新陳代謝可能發生嚴重變化，因此物質對某一替代終點的影響，在一位營養不良的人和一位營養良好的人之間可能存在很大差異（Shils et al., 2006）。在那些物質攝入量普遍不足的國家和地區從事的研究，就不能得出科學結論，因為對物質攝入量的反應可能源於對營養不良的糾正，而這種結果並不是健康聲明的目的。

另外，如果研究是在發病原因與美國顯著不同的國家或地區進行，也不能得出結論。例如，日本的主要胃癌風險因素（鹽攝入量高和幽門螺桿菌感染）顯著超過美

國。因此，就物質和胃癌關係對日本人口進行的研究資料，就不適合用於對美國人口所受的潛在影響得出結論。

觀察研究

- *收集何種資訊？*只有在證據顯示物質攝入量和生物樣本中的物質水準（或物質的新陳代謝）之間存在很強的劑量反應關係時，才應當把生物樣本（如血、尿、組織或頭髮）用於建立物質攝入量。應當有證據表明物質攝入量和生物樣本中的物質或新陳代謝水準之間（例如硒攝入量和血清硒濃度之間）存在很強的相關性⁽²³⁾。如果一個特定的生物樣本的相關度低，則不能從把該生物樣本用作攝入量生物標記的研究中得出科學結論。病例對照研究中的生物樣本不應當用於確定物質攝入量，因為物質的新陳代謝或濃度可能由於受試者患病而改變。

- *科學上可以接受並經過驗證的食品評估方法是否用於評估物質攝入量？*通常，一次 24 小時的食用量回憶或飲食記錄，不被視作評估個人通常物質攝入量的健全方法，雖然這種評估可能對確定一組受試者的平均攝入量有用處。建立飲食歷史要求與受試者進行廣泛面談。但是，飲食歷史通常也不足以用來評估物質攝入量，因為被調查者必須就通常食品及食用的份量作出判斷。食用頻率問卷包含一組數目有限的食品，如果物質的主要來源沒有包括在問卷中，就不足以評估物質攝入量。食用頻率問卷也並不總是反映某一食品的不同類別或不同的烹調方法。由於這些限制，食用頻率問卷方法的驗證對於評估食品攝入量至關重要，只有經過驗證才能從科學資料中得出結論，因為缺乏驗證可能導致飲食因素和疾病或疾病相關標記之間出現假相關⁽²⁴⁾。

- *觀察研究是否評估疾病和一種食品或食品成份之間的關係？*由於觀察研究是採用記錄食品攝入量的方法（例如食用頻率問卷、食品回憶或飲食記錄）來估算一種完整食品的攝入量，因此觀察研究的一種常見弱點是很難確知研究群體的實際物質攝入量。另外，如果物質是一種食品成份，而不是一種完整食品，則需要對每一種食品中存在的食品成份含量進行另外的估算。食品成份的含量可能基於多種因素變化，例如土壤構成、食品加工/烹調程序或儲存（時間長短、溫度）。因此，很難基於完整食品攝入量報告來確定所食用的食品成份的準確數量。

另外，包括數種食品成份（例如多種營養品補充劑）的完整食品或產品，不僅包含作為聲明主題的食品成份，還包含可能與有關食品成份新陳代謝或疾病發病原因或其他健康狀況相關的其他食品成份。由於完整食品和產品（例如多種營養品補充劑）包含許多食品成份，很難對食品成份進行孤立研究（Sempos et al., 1999）。至於基於記錄完整食品或多個食品成份攝入量的研究，不可能準確地確定作為聲明主題的食品成份對疾病風險所產生的任何觀察到的影響與下列原因有關：（1）僅僅是該食品成份；（2）與其他食品成份的交互作用；（3）獨立或共同起作用的其他食品成份；或（4）由於增加攝入相關食品成份含量豐富的食物而造成食品中包含的其他物質攝入量減少（請參見 Sempos et al. (1999)、Willett (1990)和 Willett(1998)關於確定某一食品中特定食品成份和疾病之間的關係的複雜性）。

事實上，有證據表明，在多種情形下，基於記錄常規食品攝入量的觀察研究可能顯示某種營養素對某種疾病的益處，但是後來進行的一項干預研究表明，包含營養素的食品補充劑對疾病並無益處，甚至還增加了疾病風險（Lichtenstein and Russell, 2005）。例如，以往的研究報告說beta胡蘿蔔素含量高的水果和蔬菜與降低肺癌風險有關（Peto et al., 1981）。但是，在後來的干預研究中——Alpha維他命E和Beta胡蘿蔔素干預研究（ATBC）以及胡蘿蔔素與維他命A影響試驗（CARET）——顯示Beta胡蘿蔔素補充劑增加了吸煙者和接觸石棉的工人的肺癌風險（Alpha維他命E和Beta胡蘿蔔素癌症預防研究組，1994；Omenn et al., 1996）。這些研究表明，作為食品補充劑提供的一種營養素在與許多其他食品成份共同食用時會產生不同的健康影響。另外，這些研究表明僅依賴流行病學研究的結果可能為公眾健康帶來風險，因為在此類研究中對一種營養素的測量方法是僅僅記錄常規食品的攝入量，並且由此認定該營養素和疾病風險之間的關係；實際上這種影響可能有害。出於上述原因，無法從觀察研究中得出一種食品成份和一種疾病之間的關係的科學結論。但是，觀察研究可用於測量一種完整食品和一種疾病之間的關聯。[\(25\)](#)

E、評估研究方法品質

至於在前面的評估中沒有被排除的研究，FDA 打算獨立評估每一項此類研究的方法品質。研究品質的評級分為高、中等或低。FDA 打算基於數種因素進行品質評級，這些因素涉及研究設計、資料收集、統計分析品質、測量的結果類型、除與美

國人口的相關性之外的研究人口特徵（例如選擇偏差以及提供重要的受試者資訊 [例如年齡、吸煙者]）。如果科學研究充份考慮到上述所有因素或大多數因素，FDA 就會在方法品質上給予高分。FDA 打算基於品質因素中的不足或不確定性的程度給予中等品質或低品質評分。如果研究品質很差，得到低品質評分，就不能據以得出物質/疾病關係的科學結論，這些研究將不進行進一步審查。

對於到目前為止根據本指南所述科學評估程序評估後依然保留的各項研究，FDA 用以評估其研究方法品質的因素舉例如下：

干預研究

- *研究是否進行隨機抽樣？是否具有盲性？是否提供安慰劑？*適當的隨機抽樣能夠消除可能影響研究結果的，除研究物質之外的其他內在及/或外在因素。當終點可能受到受試者意識的影響時（受試者瞭解他或她在服用某種可能有益的物質），盲性就特別重要。當結果測量是認知表現、心理狀態（例如記憶、抑鬱）或行為時，盲性至關重要。在補充劑試驗中使用安慰劑可防止受試者瞭解他/她是否在服用試驗物質。
- *是否提供了入選/排除標準以及關於研究人口特徵的重要資訊？*例如，健康或高風險受試者是否獲准在研究期間服用可能影響作為聲明主題的疾病的藥物？如果是這樣，對照組和干預組之間服藥的受試者比例是否類似？
- *對受試者減少（在結束前退出研究）是否評估？是否在報告研究的文章中加以說明？是否合理？*如果有過多受試者退出研究，則必須瞭解退出原因以及退出對干預組和安慰組人數及構成的影響。
- *如何驗證遵守了研究規程？*干預研究應當包括一種機制，核查受試者是否遵守了研究規程。例如，一項補充劑試驗應當有一種機制確定受試者多長時間服用一次補充劑。需要重點瞭解：1）受試者是否服用了研究提供的所有補充劑還是僅僅服用了其中一部分，以及2）每一組受試者中服用計量不足的人所佔的比例。
- *在進行統計分析時，使用的基線資料包含最初招收的所有受試者，還是僅僅包含完成研究的受試者？*如果有許多受試者退出，對干預組和安慰組的構成產生了不同

的影響，則必須確定就基線資料進行的統計分析是包含最初招收的所有受試者，還是僅僅包含完成研究的受試者。

- *研究是測量發病率還是測量疾病風險替代終點？*雖然疾病風險替代終點得到驗證，其結果依然不如測量實際發病率準確。觀察研究也會考慮這個品質問題。
- *如何確定發病開始時間？*當發病率是被測量的終點時，作為聲明主題的疾病必須透過醫療記錄及/或病理報告確認。依賴較不具體的記錄（例如死亡證書）是不夠的。觀察研究也會考慮這個品質問題。

觀察研究

- *對於疾病風險混雜因素是否做了足夠的調整？*一種物質/疾病關係的若干方面可能導致混雜因素。因此，對有關疾病的混雜因素作出調整很重要，以便觀察到的由於混雜因素產生的對疾病的影響，不至於被錯誤地歸因於研究物質。例如，一種疾病可能有多種與食品無關的風險因素（例如吸煙、身體品質指數和年齡對高血壓的影響）。因此，在評估鈉和血壓的關係時，要基於年齡、吸煙、身體品質指數等對風險分析作出調整。
- *在估算食品攝入量時使用了哪一類食品評估方法？*在估算「通常」食品攝入量時，經驗證的食用頻率問卷的可靠性高於飲食記錄或 24 小時回憶方法。請參見三、B 部分。

F、評估科學證據的完整性

根據本指南描述的方法，目前FDA打算評估可以從中得出科學結論的研究結果，並評定作為公開資訊的全部證據的強度。FDA打算在做這項評估時考量下列因素：研究類型（例如干預、預期隊列、病例對照、橫斷面）、先前評定的方法品質級別、各類研究的數目和樣本大小、科學證據與美國人口或目標群體的相關性、支持擬議的聲明的研究結果是否曾經複製⁽²⁶⁾、以及全部證據的總體一致性⁽²⁷⁾、⁽²⁸⁾。基於科學證據的完整性，FDA確定此類證據是否符合SSA標準，或此類證據的可信度是否足以支持關於物質/疾病關係的合格健康聲明。

在每一類研究中對下列因素進行審查：

- 研究數目和每組受試者人數
- 方法品質(高、中等或低)。
- 每一研究類型內部的研究結果(有益影響、無影響、不利影響)。要證明一項干預研究的結果有效，干預組應當與對照組在統計上顯著不同($P < 0.05$)。至於觀察研究，當數值小於或大於「1」時，風險置信區間(CI)具有顯著性。許多研究旨在分析物質和疾病線性關係(P 指趨勢)的統計顯著性。雖然該趨勢可能具有顯著性($P < 0.05$)，不同攝入量水準上(例如攝入量的三分位、四分位或五分位)⁽²⁹⁾受試者之間的差異不一定具有顯著性。在這種情形下，研究顯示為沒有影響。對影響規模(例如LDL膽固醇百分比下降)的評估對於比較一項研究內部的影響可能有用處(例如物質的兩種形式的相對影響或兩種食用頻率的相對影響)。
- 總體而言，各項研究之間在顯示有益關係方面的一致性越大，物質/疾病關係的置信水準越高。相互矛盾的結果並不排除一種關聯(因為研究設計的要素可能就是負面結果的研究中缺乏影響的原因)，但往往會削弱關聯強度的置信水準。有益影響的規模越大，就越可能存在關聯。
- 與美國總人口的相關度。例如，

顯示有益影響的研究在何種程度上與美國總人口或特定群體(例如老人、婦女)相關？

研究是否僅僅包括生活方式獨特的受試者(例如吸煙者、素食者)？

研究是否表明提供有益影響的物質的攝入量，顯著超出美國人的通常攝入量？

在評估證據的完整性時，FDA要確定這些證據是足以支持有關全體美國人口的聲明，還是支持僅僅適用於某一群體的聲明。如果證據僅僅支持關於某一群體的聲明，這種情形在聲明中必須加以說明。如果物質在用於降低風險時必須顯著超過正常美國人的攝入量，也要在聲明中加以說明。

一般而言，干預研究為聲明中所說的影響提供最強有力的證據，無論關於同一關係的現有觀察研究的結果如何。干預研究的設計旨在避免選擇偏差，以及由於偶然因素或疾病的其他混雜因素所導致的發現 (Sempos et al., 1999)。雖然物質/疾病關係的評估經常同時涉及干預研究和觀察研究，通常觀察研究不能用於排除來自更可靠的干預研究的發現 (Sempos et al., 1999)。一項干預研究不足以排除許多觀察研究的一致發現。但是，如果數項隨機控制干預研究都顯示或不顯示物質/疾病關係，則此種結果優先於任何數量的觀察研究得出的結果 (Barton, 2005)。這是因為干預研究的設計和控制，著眼於測試物質和減少疾病風險之間的因果關係的證據，而觀察研究只能識別可能的相關性。在無數例證中——例如維他命 E 和 CVD 之間的關係以及 beta 胡蘿蔔素和肺癌之間的關係——觀察研究中發現的相關性被公佈，但是，透過使用隨機對照干預研究測試可能的相關性，都找不到支持這些關係的證據 (Lichtenstein and Russell, 2005)。

G、評估顯著的科學一致性 (SSA)

顯著的科學一致性是指某一領域中的合格專家對一種看法的一致程度。科學證據包括從十分有限到不能得出結論的證據，SSA 則接近完全一致。FDA 對 SSA 的認定代表 FDA 的最佳判斷，即合格專家是否有可能一致認為，作為擬議健康聲明主題的物質/疾病關係，能夠得到科學證據的支持。FDA 的意圖是把 SSA 標準用做一個強有力的標準，標誌著對物質/疾病關係的有效性具有很強的信心。SSA 意味著關係的有效性不太可能被新興科學和現有科學的演進所推翻，雖然關係的確切性質有可能需要進一步限定。SSA 並不要求基於完全一致和沒有爭論的科學意見。SSA 在新興科學出現很長一段時間之後發生，此時能夠根據資料和資訊作出推論，但是該科學領域內還沒有就推論的有效性達成完全一致的認識。

要使合格專家對一項聲明的有效性達成明智的意見，與聲明有關的資料和資訊必須是有關科學領域中能夠得到的資料和資訊。因此，要認定 SSA 的存在，就必須得出結論說有足夠的、公開發佈的相關科學證據顯示，在多項研究中和不同研究人員從事的研究中存在一致性。而能夠說明合格專家得以取得相關證據的常用機制，就是由同行評審的科學期刊已經出版了這些資料和資訊。如果一位專家無法接觸到所有證據，那麼其意見的價值就是有限的。

在確定是否存在顯著的科學一致性時，FDA 也考量其機構外部的合格專家的意見，前提是此類專家做過評估並且其評估結果已經公開發表。例如，FDA 打算考量：

- 根據一個具有公信力的獨立機構專門為此目的召開的「專家小組會議」的意見所編寫的文件；
- 一個聯邦政府科學機構的意見或建議，例如國家衛生研究院（NIH）、疾病控制和預防中心（CDC）或國家科學院（NAS）；
- 一個獨立的專業機構的意見，例如美國兒科醫生學會（AAP）營養委員會、美國心臟學會（AHA）、美國癌症學會（ACS）或由國家衛生研究院（NIH）召集的工作組或其他小組；
- 對二級科學文獻中的資料和資訊進行總結和批評的評論文章。

FDA 對於聯邦科學機構得出的結論給予很大信任，特別是此類機構對一種物質/疾病關係的有效性證據進行評審，認為足以向公眾提出飲食建議。當聯邦科學機構得出結論支持一種物質/疾病關係的有效性時，FDA 通常認定存在顯著的科學一致性。其他專家機構得出的結論也可能被用於支持關於 SSA 的認定。雖然作為個人的外部專家的意見在評估 SSA 時也予以考量，但是這種意見不足以支持達到標準的結論，特別是聯邦科學機構、專家小組或獨立專業機構對同樣證據審查後沒有對該專家的意見表示支持。外部專家或專家小組的意見最能發揮作用的情形是，有適當理由認為他們代表同一領域中更多合格專家的意見。最重要的是，外部專家意見的相關性取決於所分析的證據是否在若干要素上適用於聲明，例如對物質和疾病的規格說明和測量結果。

即使不存在合格專家得出的結論（例如如果支持擬議健康聲明的資料相對較新，並且尚未經過一個獨立專家小組或機構的審查），如果存在一批令人信服的相關證據，也可能促使 FDA 認定存在顯著的科學一致性。由於每一種聲明的物質/疾病關係的性質可能各有不同，有必要基於個案同時考量一致性的程度和不同意見的性質。如果要根據主觀的定量標準或僵化的標準來評估科學一致性，所導致的不靈活可能造成某些有效的聲明被駁回，在這種情形下，雖然存在不同意見，但此類意見並不令人信服。

適用顯著的科學一致性標準的意圖是：在依賴一組健全的、相關的科學資料時力求客觀；在確認支持不同的物質/疾病關係的有效性時，所需要的資料數量和類別的差異時力求靈活；在長期內隨著研究問題和試驗方法的重新界定，而在確認重新評估資料的需求方面力求適應變化。

H、合格健康聲明的語言具體性

當一種物質疾病關係的證據可信但未達到 SSA 標準時，就這種關係擬議的聲明應當包括限定性語言，說明科學證據在何種程度上支持這種關係。

健康聲明語言應當具體而準確地反映科學證據的程度。但是，科學證據中的缺口有時可能會限制可歸入聲明的資訊。例如，當科學證據有限但可信時，合格健康聲明可能無法指明降低疾病風險所需要的物質攝入量。

根據 FDA 的健康聲明規章，如果規章沒有界定何為食品中物質的「高」含量，健康聲明必須具體給出為取得所聲明的影響每天必須攝入的物質數量 (21 CFR 101.14(d)(2)(vii))。FDA 在營養成份聲明規章中把「高」含量界定為食品包含物質每天攝入量的 20% 或更多 (21 CFR 101.54(b))。因此，如果尚未建立物質的每日攝入量，FDA 就無法確定何為物質的「高」含量。如果作為聲明主題的物質沒有每日攝入量數值，FDA 會確定為達到所聲明的影響每日必須的攝入量，前提是有足夠的證明可以作出這種決定。請參見 21 CFR 101.83(c)(2)(G) (健康聲明規章：植物甾醇/甾烷醇酯和降低冠心病風險)。但是，有時關於降低風險的可信資料不夠具體，FDA 甚至不足以憑藉該資料確定普通美國人所需要的可能攝入量。請參見 FDA2004 年 9 月 8 日關於 omega-3 脂肪酸和降低冠心病風險的合格健康聲明的執行裁量書 (Martek 申請)。

如果有可信的證據表明物質和疾病之間的關係，有必要確定物質是在該關係中獨立發揮作用，還是其作用係基於包含或取代其他物質。有時，對物質的獨立作用的評估可能很困難，例如物質屬於一種常規食品或大量營養素 (例如脂肪或碳水化合物)。在評估一種常規食品或大量營養素可能產生的健康影響的研究中，把二者中任意一種包含在食品中，通常要求排除其他常規食品或大量營養素 (即取代產生相

等熱量的食品)。如果確定物質不是獨立發揮作用及/或需要減少或包含另一種物質才能顯示有益影響，聲明語言就需要說明這種情形。

I、對現有SSA或合格健康聲明的重新評估

無論是否有人提出申請，FDA 都有可能重新評估一項健康聲明，在這樣做時它打算使用上文所述的科學評估程序。為了透過健康聲明審查為公眾健康帶來最大益處，FDA 打算對新發佈的資訊進行評估，以確定基於新資訊是否有必要對現有 SSA 或合格健康聲明作出改動。例如，新的科學證據可能：(1) 支持對 SSA 或合格健康聲明的語言作出修定，(2) 支持把 SSA 聲明改為合格健康聲明 (QHC)，或支持把 QHC 改為 SSA 聲明，或(3) 對作為健康聲明主題的物質提出安全關切，或不再支持健康聲明 (SSA 或 QHC)。

四、參考資料

- 美國癌症學會，「癌症資料與圖表」，2004。
- Alpha 維他命 E、Beta 胡蘿蔔素癌症預防研究小組。「維他命 E、Beta 胡蘿蔔素對男性吸煙者肺癌和其他癌症發病率的影響」，「*新英格蘭醫學雜誌*」1994；330:1029-1035。
- Barton S.，「哪些臨床研究提供最佳證據？最佳隨機對照試驗 (RCT) 仍然優於最佳觀察研究」，「*英國醫學雜誌*」2000; 321:255-256。
- Cade J、Thompson R、Burley V、Warm D，「食品攝入頻率問卷的編製、驗證和使用——總結評估」，「*公共健康營養*」2002; 5:567-587。
- 聯邦司法中心，「科學證據參考手冊」第二版，2000。
- Flegal KM，「就食品和營養攝入的影響評估流行病學證據」，「*美國臨床營養雜誌*」1999; 69:1339S-1344S。
- Greer N、Mosser G、Logan G、Halaas GW，「關於證據評級的一種實用方法」，「*品質改進聯合委員會雜誌*」，2000; 26:700-712。
- Hill AB，「環境和疾病：關聯或因果？」，「*皇家醫學會會議記錄*」，1965;58:295-300。
- Hord NG、Fenton JI，「環境最重要：探索常態和癌症前期微環境以破解食品和癌症風險的難題」，「*分子營養和食品研究*」，2007; 51:100-106。

- IOM, 醫學研究會, 「食品補充劑: 一種安全評估框架」, 國家科學院出版社, 哥倫比亞特區華盛頓市, 2005。
- Kraemer HC、Lowe KK、Kupfer DJ, 「祝您健康: 如何理解研究得出的風險結論?」 牛津大學出版社, 2005。
- Kris-Etherton PM、Dietschy J, 「為檢查脂肪酸對心血管疾病風險因素的影響的研究設計標準: 人體和動物研究」, 「美國臨床營養雜誌」1997; 65:1590S-1596S。
- Lichtenstein AH、Russell RM, 「至關重要的營養素: 食品或補充劑?」, 「美國醫學學會雜誌」2005; 294:351-358。
- Milner JA, 「食品與癌症: 事實與爭議」, 「營養與癌症」2006; 56: 216-224。
- 國家癌症研究所, 「癌症術語詞典」, <http://www.cancer.gov/dictionary>
- 國家研究理事會, 「食品與健康: 減少慢性病風險的作用」, 國家科學院出版社, 哥倫比亞特區華盛頓市, 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 胡蘿蔔素和維他命 A 對肺癌和心血管疾病的綜合影響」, 「新英格蘭醫學雜誌」1996; 334:1150-1155。
- Peto R、Doll R、Buckley JD、Sporn MB, 「食品 beta 胡蘿蔔素能否顯著減少人體癌症發病率?」, 「自然」1981; 290:201-208。
- Sempos CT、Liu K、Earnst ND, 「食品與營養攝入: 在評估流行病學資料時應當考量什麼?」, 「美國臨床營養雜誌」1999; 69:1330S-1338S。
- Torun B, 「蛋白質能量營養失調」, 摘自「健康與疾病的現代營養」, Williams 和 Williams 出版社, 紐約 2006。
- Spilker B, 「臨床研究指南」, 雷文出版社, 紐約, 1991。
- Subar AF、Thompson FE、Kipnis V、Midthune D、Hurwitz P、McNutt S、McIntosh A、Rosenfeld S, 「對布勞克 (Block)、惠萊特 (Willett) 和國家癌症研究所食品攝入頻率問卷的比較驗證」, 「美國流行病學雜誌」2001; 154: 1089-1099。

- Szklo M、Nieto FJ，「流行病學：超越基本因素」，阿斯彭出版社，2000。
- Willett W.C，「營養流行病學概述」，「營養流行病學」，牛津大學出版社，牛津，1990。
- Willett W.C，「分析和報告營養資料中的問題」，摘自「營養流行病學」，第二版，牛津大學出版，牛津，1998。
- Wilson E.B，「科學研究簡介」，通用出版公司，多倫多，1990。

注釋：

[\(1\)](#)本指南的編寫單位是美國食品與藥物管理局食品安全與應用營養中心下屬的營養、標籤和食品補充劑辦公室。

[\(2\)](#)為敘述方便，「疾病」在本文中代表「疾病或與健康有關的狀況」，其定義是對人體器官、部位、結構或系統造成損壞，使之無法發揮正常作用（例如心血管疾病），或導致此類功能失常的一種健康狀況（例如高血壓）。21 CFR 101.14(a)(5)。

[\(3\)](#)本新指南文件用於取代FDA發佈的另一份指南，即「行業與FDA指南：過渡性基於證據的科學資料評級系統」，該指南的內容涉及對合格健康聲明的科學審查。雖然該指南包括關於評估科學證據強度的章節，但本新指南文件並不包括這樣的章節，因為目前正在研究消費者如何理解可用於描述健康聲明證據強度的各種可能的評級系統。FDA打算重新審查其評級系統，並且在這些研究完成後發佈適當指南。另外，本指南文件還取代FDA的另一份指南，即「行業指南：常規食品 and 食品補充劑健康聲明審查中的顯著的科學一致性」。

[\(4\)](#)1997年，國會制定了「食品與藥物管理現代化法」，該法建立了一種有關健康聲明的備選授權程序，該程序適用於基於某些聯邦科學機構或國家科學院作出的權威性聲明而提出的健康聲明。本指南文件並不涉及該備選程序。

(5)請參見「行業與FDA指南：常規人類食品和人類食品補充劑標籤中的合格健康聲明的過渡性授權程序」（附件A），2003年7月10日 (<http://www.cfsan.fda.gov/~dms/hclmgi3.html>)

(6)請參見 21 U.S.C. 321(ff)(1)。

(7)在 21 CFR 101.14(a)(3)中，「營養價值」指透過促進生長、補充重要營養素或提供能量等程序維持人類生命的價值。

(8)混雜因素是指與有關疾病和干預相關的因素，如果不進行控制，這些因素會妨礙研究人員得出一種干預能夠產生結果的結論。

(9)自由生活的群體指根據自己的意願選擇食品和生活方式（例如吸煙、飲酒和健身）的群體。

(10)攝入量生物標記是對生物樣本中的物質新陳代謝物（例如血清硒）或物質本身的測量，這些測量數字經過驗證，能夠證實該物質的攝入量。

(11)偏差是一類系統性錯誤，可能源於受試者選擇缺陷（選擇偏差），也可能源於暴露和疾病結果測量缺陷（資訊偏差）（Szklo 和 Nieto, 2000）。

(12)相對風險指有暴露的人的風險（即發病率）和無暴露的人的風險之比例。在可能使用的隊列中，其計算方法是測量有病和沒有病的受試者的暴露程度。經調整的相對風險對潛在混雜因素加以控制。

(13)例如，病例對照研究可以是評估與人口中疾病發病率和分佈狀況有關的參數的研究，例如主要死亡原因。

(14)發病概率指有暴露和沒有暴露的人的發病率對比。在病例對照研究中，其計算方法是基於對物質的暴露測量受試者的疾病發生率。經調整發病概率對潛在混雜因素加以控制。

(15)少數橫斷面研究是時間系列研究，比較不同階段中的結果（例如，在一個五年階段內某一具體結果的發生率，在接下來另一個五年時間內是否有變化）。

(16) 相關係數(r)測量兩個變量之間的互相依賴性，例如一種物質的攝入量和一種疾病的發病率。其表達方式為 - 1 至 +1，- 1 表示完美的負相關，0 表示不相關，+1 表示完美的正相關。請參見「Webster's II New Riverside University Dictionary (Riverside Publishing Co., 1984)」。因此，相關係數越接近量表中的終點之一，兩個變量之間的關係越密切。

(17) 評論文章總結就某一題目進行的具體研究的結果。

(18) 其他例子包括專書文章、摘要、信函和委員會報告。

(19) 綜合分析是系統地綜合和評估已經完成或終止的臨床試驗的結果的程序 (Spilker, 1991)。

(20) 風險生物標記是表示與疾病風險相關的生理狀態變化的生物標記。

(21) 在隨機抽樣前向所有研究組所提供的食品。

(22) 交叉研究中的時間階段，在該階段內受試者不接受干預。

(23) 相關性透過相關係數(r)評估。相關係數從 - 1 (負相關) 到 +1 (正相關)。越接近 1，相關度越高；越接近 0，相關度越低。

(24) 「食品攝入頻率問卷方法的驗證至關重要，因為不正確的資訊可能導致飲食因素與疾病或疾病相關標記之間的錯誤關聯。」Cade, J、Thompson, R、Burley, V、和Warm D，「食品攝入頻率問卷的編製、驗證和使用——總結評估」，「公共健康營養」，5：573 頁，2002 年。另請參見Subar, A., et al., 「對布勞克 (Block)、惠萊特 (Willett) 和國家癌症研究所食品攝入頻率問卷的比較驗證」，「美國流行病學雜誌」，154: 1089-1099, 2001。

(25) 在 *Pearson v. Shalala* 一案中，哥倫比亞特區巡迴法院指出，FDA已經「從邏輯上確定」服用食品補充劑中的抗氧化維他命，沒有科學證據表明能夠降低癌症風險，現有研究僅考察了包含抗氧化維他命的食品，而這些食品對減輕癌症風險的影響可能源於它們包含的其他物質。164 F.3d 650, 568 (D.C. Cir. 1999)。但是，

哥倫比亞特區巡迴法院的結論是，雖然FDA對授予抗氧化維他命合格健康聲明存有顧慮，但該問題可以透過同時發佈一條醒目的免責聲明來解決，免責聲明可指出關於該聲明的證據並不是定論，因為支持該聲明的研究所基於的食品包含其他物質，而對降低癌症風險產生作用的可能實際上是這些其他物質。(同上)。法院指出，FDA並沒有聲稱有關食品補充劑會「威脅消費者的健康與安全。」(同上)，656頁。但是，基於對包含這些營養素的食品的研究，要允許基於個別營養素提出合格健康聲明還有一個更具有根本性的問題——而不僅僅是哥倫比亞巡迴法院認為透過提供一條免責聲明就可以解決的問題——即使可以有把握的確定食品中具體成份的影響，近期關於營養 - 食品交互反應以及食品、生物參數和疾病之間的關係具有複雜性的一些科學發現顯示，在吃一種或多種食品時可能產生的健康益處，在以食品補充劑的形式攝入或透過其他食品攝入時不一定會產生。請參見Lichtenstein和Russell (2005)。例如，關於食品補充劑的研究不僅證實，透過食品攝入某些營養素有關的益處在營養素作為補充劑攝入時不一定出現，而且其中一些研究實際上表明，這些營養劑增大了它們原本用於防止的疾病的風險。(同上)。因此，基於對某一具體食品或某些食品的攝入的研究所提供的資訊，不足以支持營養素本身能夠產生有益影響的科學結論。另外，即使營養素是透過其他食品攝入，而不是以補充劑形式攝入，生理學的影響也可能不一樣，因為食品組合可能影響營養素的生物利用度和生物活性。(同上)。

[\(26\)](#)科學結果的複製對評估科學證據的強度很重要 (Wilson, E.B, 「科學研究簡介」, 多佛出版社, 1990; 46-48 頁)。

[\(27\)](#)在本指南中，「一致性」的定義是指從中得出關於物質/疾病關係的科學結論的各項研究之間的吻合程度。

[\(28\)](#)類似研究和不同研究設計的結果之一致性，對於評估科學證據的因果關係和強度很重要 (Hill A.B, 「環境與疾病：關聯或因果？」*皇家醫學會會議記錄* 1965;58:295-300)；另請參見「評估科學證據的系統」，美國醫療保健研究與質量管理署，該文章把「一致性」定義為「使用類似和不同研究設計所報告的結果的類似程度。」 [<http://www.ahrq.gov/clinic/epcsums/strengthsum.htm#Contents>]

(29)攝入量的三分位、四分位和五分位是把研究人口劃分為三組、四組或五組，以至物質的平均攝入量在不同組別間形成差異（例如攝入量最低的組代表攝入量最低的三分之一，攝入量最高的組代表最高的三分之一）。在劃分研究群體時使每一組擁有同樣數目的受試者。

*更新的參照網頁：

- [對 2004 年 1 月 27 日健康聲明申請的答復：綠茶和降低癌症風險健康聲明](#)
2005 年 6 月 30 日
 - [關於「西紅柿和前列腺、卵巢、胃和胰腺癌 \(American Longevity申請 \)」的信函](#)2005 年 11 月 8 日
 - [關於鈣和結腸/直腸、乳房和前列腺癌以及復發性結腸息肉的信函](#) 2005 年 10 月 12 日
-

本文件取代 2007 年發佈的[草稿](#)。