

「機能性表示食品」制度における機能性に関する科学的根拠の
検証一届け出られた研究レビューの質に関する検証事業
報告書
(ダイジェスト版)

平成28年3月

消費者庁

本報告書は、消費者庁の委託を受け、みずほ情報総研株式会社が
有識者によるワーキンググループを設置し、取りまとめたものである。

「機能性表示食品」制度における機能性に関する科学的根拠の検証
―届け出られた研究レビューの検証事業

ワーキンググループ 委員一覧

大 室 弘 美 武蔵野大学薬学部 教授

○折 笠 秀 樹 富山大学バイオ統計学・臨床疫学講座 教授

◎上 岡 洋 晴 東京農業大学地域環境科学部 教授

北湯口 純 雲南市立身体教育医学研究所うなん 主任研究員

佐 山 暁 子 聖路加国際大学学術情報センター図書館 司書

島 田 美樹子 日高リハビリテーション病院栄養課 管理栄養士

唐 文 涛 東京大学大学院薬学系研究科博士課程

眞喜志 ま り 東邦大学習志野メディアセンター 司書

吉 崎 貴 大 東洋大学食環境科学部 助教

◎委員長 ○副委員長

(50音順、敬称略)

第1章 制度の概要と検証事業の目的

平成 26 (2014) 年 7 月 30 日に消費者庁が「食品の新たな機能性表示制度に関する検討会報告書」を公表し、平成 27 (2015) 年 4 月から「機能性表示食品」制度（以下「本制度」という。）が始まった。

本制度は食品表示法に基づく食品表示基準に規定されており、消費者庁長官に届出を行うことで、安全性及び機能性に関する一定の科学的根拠に基づき、食品関連事業者（食品表示法第 2 条第 3 項第 1 号）の責任において、健康の維持及び増進に資する特定の保健の目的（疾病リスクの低減に係るものを除く。）が期待できる旨の表示を行うことが可能となる制度である。従来、食品の機能性表示は特定保健用食品と栄養機能食品においてのみ可能であったが、本制度の施行により、機能性を分かりやすく表示した商品の選択肢が増え、消費者がそうした商品の情報を得て選択できることが期待されている。

機能性に関する科学的根拠は、「機能性表示食品の届出等に関するガイドライン」(平成 27(2015)年 3 月 30 日公表)^{注1}（以下「ガイドライン」という。）にのっとり「最終製品を用いた臨床試験」、「最終製品又は機能性関与成分に関する研究レビュー」のいずれかで示すこととされている。平成 27 (2015) 年 10 月 31 日時点では、122 件の食品の届出が消費者庁ウェブサイト公表（うち 2 件は撤回）されている。

研究レビューは、研究論文等の文献から得られる知見を体系的に整理したもので、一般的にシステマティック・レビュー（以下「SR」という。）と呼ばれる。本制度における機能性に関する研究レビューの方法については、ガイドラインに「最終製品又は機能性関与成分に関する研究レビューの実施に当たっての留意事項」が明記されている。食品関連事業者（以下「届出者」という。）は定性的研究レビュー又は定量的研究レビュー（メタアナリシス）を実施し、「totality of evidence」（関連研究について、肯定的・否定的内容及び研究デザインを問わず検討し、総合的観点から肯定的といえるかを判断）の観点から、表示しようとする機能性について肯定的と判断できるものに限り、科学的根拠になり得るとされている。

本事業では、本制度において届出された食品について、その機能性の科学的根拠として提出された研究レビュー（以下「届出 SR」という。）について科学的な評価を行い、平成 28 年度以降の本制度の適正な運用に向けた課題の抽出、届出 SR の質を高める方策等の検討を行うことを目的として、下記の業務を実施した。

1. PRISMA 声明チェックリストに基づく検証（第 3 章）
2. 本制度における届出 SR に特有の報告方法の質の検証（第 4 章）
3. 機能性表示食品の届出 SR の改善のポイントの提示（第 5 章、付録）

^{注1} 「機能性表示食品の届出等に関するガイドライン」消費者庁 2015.3.30. <http://www.caa.go.jp/foods/pdf/150330_guideline.pdf>（最終閲覧 2016 年 3 月 22 日）

第2章 報告書で取り扱う範囲と報告の位置付け

本事業では、PRISMA 声明チェックリスト（2009 年）に基づく評価のほか、多くの評価方法を検討し、届出 SR の評価を行った。そのため、本報告書の取り扱う範囲や、内容の正確な理解と解釈を促す上で注意すべき事項をあらかじめ明確にする必要がある。そこで本章では、本事業の検証対象とした範囲と、報告の位置付けについて示す。

本事業の目的は、届出 SR が正しい方法論に基づいて実施されたかを明らかにし、届出 SR の質を高めるための提案を行うことにある。具体的には、記述すべき事項などに関して、不履行や記入漏れ、不確かな点がないかを検証し、どのように記述すべきかを提案することである。したがって、届出 SR で記述された本文と図表が本事業の対象である。

食品の機能性表示における主な関心事は、機能性関与成分や最終製品は、「本当に安全なのか」、「作用機序は本当に明確なのか」、「本当に機能があるのか」の 3 点であると考えられる。しかし、これらを明らかにするには、届出 SR や別の資料で提出された「引用・参考にした一次研究」に遡り、個々の研究報告の質を評価しなければ結論を見出すことはできない。

それはまた別の次元の検証研究であることから、本事業では、機能性関与成分や最終製品における安全性、作用機序、機能性の有無をカバーするものではなく、前述のとおり報告方法の質の評価に限定している。

第3章 PRISMA 声明チェックリストに基づく検証

第1項 「PRISMA 声明チェックリスト：機能性表示食品のための拡張版」に基づいた届出 SR の検証結果

1.方法

平成 27 (2015) 年 10 月 31 日までに公表された食品 122 件から「届出の撤回 (2 件)」、「最終製品を用いた臨床試験による届出 (21 件)」を除いた 99 件の食品の届出で提出された届出 SR (99 編) のうち、既に届出済みのものと重複する届出 SR (48 編) を除外した結果、51 編の届出 SR が質評価の対象となった (図 1)。なお、1 つの届出に複数の機能性を表示し、それぞれについて届出 SR が提出されていた場合は、それらをまとめて 1 編の届出 SR として扱った。

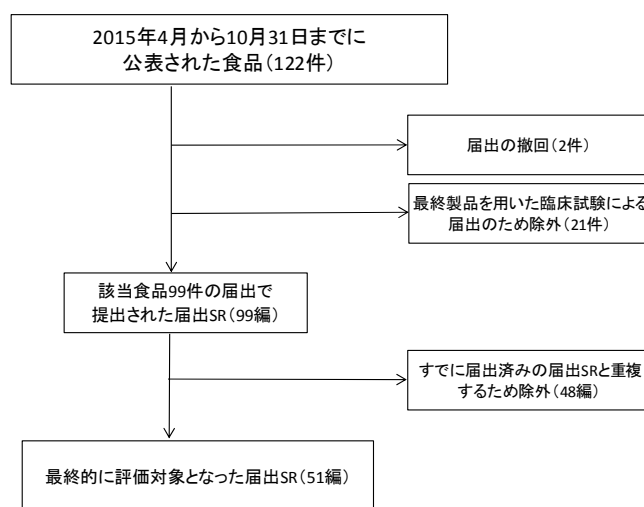


図 1 対象とする届出 SR の選別過程

なお、対象とした届出 SR 及び、除外した届出 SR と除外理由については報告書別添資料に一覧を掲載した。

質評価に当たっては、27 の大項目からなる「PRISMA 声明チェックリスト (2009 年)」に合計 45 の下位項目を設定し、「PRISMA 声明チェックリスト：機能性表示食品のための拡張版」を作成した。

2.結果

「不履行・記述なし」や「不十分又は不明瞭な記述」を記述の不備とみなし、不備があった届出 SR の編数 (度数) と率 (不備率) を算出した。その結果を、評価対象とした届出 SR 全体については表 1 に、うち、メタアナリシスを行った届出 SR については表 2 に、メタアナリシスを行っていない届出 SR については表 3 にそれぞれまとめた。

不備率の高い項目については特に注意して適正な記述を行う必要がある。各項目の評価結果から得られた改善すべき課題点、具体的な記述方法は、報告書本文を参照していただきたい。

表1 「PRISMA 声明チェックリスト:機能性表示食品のための拡張版」に基づく質評価の結果一覧
【届出 SR(n=51)】

項目番号	「PRISMA 声明チェックリスト:機能性表示食品のための拡張版」における評価項目	届出 SR(n=51)		
		「不十分又は不明瞭な記述」度数(率)	「不履行・記述なし」度数(率)	不備数(率)
#1	タイトル	0 (0.0%)	0 (0.0%)	0 (0.0%)
#2	構造化抄録	33 (64.7%)	0 (0.0%)	33 (64.7%)
#3	論拠	21 (41.2%)	4 (7.8%)	25 (49.0%)
#4a	参加者の記述	4 (7.8%)	2 (3.9%)	6 (11.8%)
#4b	介入の記述	4 (7.8%)	2 (3.9%)	6 (11.8%)
#4c	比較の記述	3 (5.9%)	5 (9.8%)	8 (15.7%)
#4d	アウトカムの記述	4 (7.8%)	2 (3.9%)	6 (11.8%)
#5a	レビュー・プロトコルの有無	9 (17.6%)	17 (33.3%)	26 (51.0%)
#5b	プロトコルへのアクセスの可否	0 (0.0%)	49 (96.1%)	49 (96.1%)
#5c	プロトコルの web address、登録番号の有無	0 (0.0%)	49 (96.1%)	49 (96.1%)
#6a	研究の特性の記述	16 (31.4%)	1 (2.0%)	17 (33.3%)
#6b	報告の特性の記述	19 (37.3%)	3 (5.9%)	22 (43.1%)
#7	情報源	20 (39.2%)	2 (3.9%)	22 (43.1%)
#8	検索	7 (13.7%)	1 (2.0%)	8 (15.7%)
#9a	スクリーニング方法に関する記述	21 (41.2%)	0 (0.0%)	21 (41.2%)
#9b	適格性に関する記述	12 (23.5%)	0 (0.0%)	12 (23.5%)
#9c	採択基準に関する記述	17 (33.3%)	0 (0.0%)	17 (33.3%)
#10	データの収集プロセス	17 (33.3%)	7 (13.7%)	24 (47.1%)
#11	データ項目	12 (23.5%)	1 (2.0%)	13 (25.5%)
#12a	バイアス・リスク	27 (52.9%)	2 (3.9%)	29 (56.9%)
#12b	非直接性	20 (39.2%)	6 (11.8%)	26 (51.0%)
#12c	不精確	12 (25.5%)	4 (8.5%)	16 (34.0%)
#12d	非一貫性	19 (37.3%)	3 (5.9%)	22 (43.1%)
#13	要約尺度	9 (21.4%)	7 (16.7%)	16 (38.1%)
#14a	研究結果の統合方法の記述	2 (22.2%)	2 (22.2%)	4 (44.4%)
#14b	一致性の尺度の記述	1 (11.1%)	1 (11.1%)	2 (22.2%)
#15a	臨床試験登録の検索	2 (3.9%)	38 (74.5%)	40 (78.4%)
#15b	著者への問合せ	1 (2.0%)	46 (90.2%)	47 (92.2%)
#15c	(事後メタアナリシス時)ファンネルプロット	0 (0.0%)	1 (11.1%)	1 (11.1%)
#15d	研究内での選択的報告及びその他の記述	15 (31.9%)	9 (19.1%)	24 (51.1%)
#16	追加的解析	0 (0.0%)	3 (37.5%)	3 (37.5%)
#17	研究の選択	1 (2.0%)	0 (0.0%)	1 (2.0%)
#18	研究の特性	10 (19.6%)	0 (0.0%)	10 (19.6%)
#19	研究内のバイアス・リスク	13 (25.5%)	1 (2.0%)	14 (27.5%)
#20a	各介入群の単純な要約データの記述	19 (38.8%)	5 (10.2%)	24 (49.0%)
#20b	効果の推定量と信頼区間の記述	1 (11.1%)	1 (11.1%)	2 (22.2%)
#21	結果の統合	1 (12.5%)	0 (0.0%)	1 (12.5%)
#22	全研究のバイアス・リスク	17 (33.3%)	0 (0.0%)	17 (33.3%)
#23	追加的解析	1 (11.1%)	4 (44.4%)	5 (55.6%)
#24	エビデンスの要約	31 (64.6%)	8 (16.7%)	39 (81.3%)
#25a	研究レベルとアウトカムレベルでの限界の記述	18 (35.3%)	15 (29.4%)	33 (64.7%)
#25b	レビューレベルでの限界の記述	20 (39.2%)	10 (19.6%)	30 (58.8%)
#26	結論	29 (56.9%)	10 (19.6%)	39 (76.5%)
#27a	SR の資金源と、その他の支援	8 (15.7%)	7 (13.7%)	15 (29.4%)
#27b	SR における資金提供者の役割	3 (6.1%)	9 (18.4%)	12 (24.5%)

※不備数(率) : 「不履行・記述なし」と「不十分又は不明瞭な記述」の合計数(率)。不備率は届出 SR 計 51 編から非該当を除いたものを分母として算出。

※#14a、#14b、#15c、#16、#20b、#21、#23 は、メタアナリシスを実施した届出 SR (n=9) についてのみ算出。

表2 「PRISMA 声明チェックリスト:機能性表示食品のための拡張版」に基づく質評価の結果一覧
【メタアナリシスを実施した届出 SR(n=9)】

項目番号	「PRISMA 声明チェックリスト:機能性表示食品のための拡張版」における評価項目	メタアナリシスを実施した届出 SR(n=9)		
		「不十分又は不明瞭な記述」度数(率)	「不履行・記述なし」度数(率)	不備数(率)
#1	タイトル	0 (0.0%)	0 (0.0%)	0 (0.0%)
#2	構造化抄録	6 (66.7%)	0 (0.0%)	6 (66.7%)
#3	論拠	3 (33.3%)	2 (22.2%)	5 (55.6%)
#4a	参加者の記述	0 (0.0%)	1 (11.1%)	1 (11.1%)
#4b	介入の記述	0 (0.0%)	1 (11.1%)	1 (11.1%)
#4c	比較の記述	1 (11.1%)	1 (11.1%)	2 (22.2%)
#4d	アウトカムの記述	0 (0.0%)	0 (0.0%)	0 (0.0%)
#5a	レビュー・プロトコルの有無	1 (11.1%)	4 (44.4%)	5 (55.6%)
#5b	プロトコルへのアクセスの可否	0 (0.0%)	8 (88.9%)	8 (88.9%)
#5c	プロトコルの web address、登録番号の有無	0 (0.0%)	8 (88.9%)	8 (88.9%)
#6a	研究の特性の記述	2 (22.2%)	1 (11.1%)	3 (33.3%)
#6b	報告の特性の記述	1 (11.1%)	1 (11.1%)	2 (22.2%)
#7	情報源	3 (33.3%)	0 (0.0%)	3 (33.3%)
#8	検索	1 (11.1%)	0 (0.0%)	1 (11.1%)
#9a	スクリーニング方法に関する記述	4 (44.4%)	0 (0.0%)	4 (44.4%)
#9b	適格性に関する記述	1 (11.1%)	0 (0.0%)	1 (11.1%)
#9c	採択基準に関する記述	2 (22.2%)	0 (0.0%)	2 (22.2%)
#10	データの収集プロセス	6 (66.7%)	0 (0.0%)	6 (66.7%)
#11	データ項目	0 (0.0%)	1 (11.1%)	1 (11.1%)
#12a	バイアス・リスク	3 (33.3%)	2 (22.2%)	5 (55.6%)
#12b	非直接性	2 (22.2%)	3 (33.3%)	5 (55.6%)
#12c	不精確	1 (11.1%)	1 (11.1%)	2 (22.2%)
#12d	非一貫性	1 (11.1%)	1 (11.1%)	2 (22.2%)
#13	要約尺度	1 (12.5%)	1 (12.5%)	2 (25.0%)
#14a	研究結果の統合方法の記述	2 (22.2%)	2 (22.2%)	4 (44.4%)
#14b	一致性の尺度の記述	1 (11.1%)	1 (11.1%)	2 (22.2%)
#15a	臨床試験登録の検索	0 (0.0%)	8 (88.9%)	8 (88.9%)
#15b	著者への問合せ	0 (0.0%)	7 (77.8%)	7 (77.8%)
#15c	(事後メタアナリシス時)ファンネルプロット	0 (0.0%)	1 (11.1%)	1 (11.1%)
#15d	研究内での選択的報告及びその他の記述	2 (22.2%)	1 (11.1%)	3 (33.3%)
#16	追加的解析	0 (0.0%)	3 (37.5%)	3 (37.5%)
#17	研究の選択	1 (11.1%)	0 (0.0%)	1 (11.1%)
#18	研究の特性	3 (33.3%)	0 (0.0%)	3 (33.3%)
#19	研究内のバイアス・リスク	2 (22.2%)	0 (0.0%)	2 (22.2%)
#20a	各介入群の単純な要約データの記述	1 (11.1%)	0 (0.0%)	1 (11.1%)
#20b	効果の推定量と信頼区間の記述	1 (11.1%)	1 (11.1%)	2 (22.2%)
#21	結果の統合	1 (12.5%)	0 (0.0%)	1 (12.5%)
#22	全研究のバイアス・リスク	3 (33.3%)	0 (0.0%)	3 (33.3%)
#23	追加的解析	1 (11.1%)	4 (44.4%)	5 (55.6%)
#24	エビデンスの要約	6 (66.7%)	1 (11.1%)	7 (77.8%)
#25a	研究レベルとアウトカムレベルでの限界の記述	2 (22.2%)	2 (22.2%)	4 (44.4%)
#25b	レビューレベルでの限界の記述	3 (33.3%)	2 (22.2%)	5 (55.6%)
#26	結論	3 (33.3%)	4 (44.4%)	7 (77.8%)
#27a	SR の資金源と、その他の支援	0 (0.0%)	2 (22.2%)	2 (22.2%)
#27b	SR における資金提供者の役割	0 (0.0%)	1 (11.1%)	1 (11.1%)

※不備数(率) : 「不履行・記述なし」と「不十分又は不明瞭な記述」の合計数(率)。不備率は届出 SR 計 51 編から非該当を除いたものを分母として算出。

表3 「PRISMA 声明チェックリスト:機能性表示食品のための拡張版」に基づく質評価の結果一覧
【メタアナリシスを実施していない届出 SR(n=42)】

項目番号	「PRISMA 声明チェックリスト:機能性表示食品のための拡張版」における評価項目	メタアナリシスを実施していない届出 SR(n=42)		
		「不十分又は不明瞭な記述」度数(率)	「不履行・記述なし」度数(率)	不備数(率)
#1	タイトル	0 (0.0%)	0 (0.0%)	0 (0.0%)
#2	構造化抄録	27 (64.3%)	0 (0.0%)	27 (64.3%)
#3	論拠	18 (42.9%)	2 (4.8%)	20 (47.6%)
#4a	参加者の記述	4 (9.5%)	1 (2.4%)	5 (11.9%)
#4b	介入の記述	4 (9.5%)	1 (2.4%)	5 (11.9%)
#4c	比較の記述	2 (4.8%)	4 (9.5%)	6 (14.3%)
#4d	アウトカムの記述	4 (9.5%)	2 (4.8%)	6 (14.3%)
#5a	レビュー・プロトコルの有無	8 (19.0%)	13 (31.0%)	21 (50.0%)
#5b	プロトコルへのアクセスの可否	0 (0.0%)	41 (97.6%)	41 (97.6%)
#5c	プロトコルの web address、登録番号の有無	0 (0.0%)	41 (97.6%)	41 (97.6%)
#6a	研究の特性の記述	14 (33.3%)	0 (0.0%)	14 (33.3%)
#6b	報告の特性の記述	18 (42.9%)	2 (4.8%)	20 (47.6%)
#7	情報源	17 (40.5%)	2 (4.8%)	19 (45.2%)
#8	検索	6 (14.3%)	1 (2.4%)	7 (16.7%)
#9a	スクリーニング方法に関する記述	17 (40.5%)	0 (0.0%)	17 (40.5%)
#9b	適格性に関する記述	11 (26.2%)	0 (0.0%)	11 (26.2%)
#9c	採択基準に関する記述	15 (35.7%)	0 (0.0%)	15 (35.7%)
#10	データの収集プロセス	11 (26.2%)	7 (16.7%)	18 (42.9%)
#11	データ項目	12 (28.6%)	0 (0.0%)	12 (28.6%)
#12a	バイアス・リスク	24 (57.1%)	0 (0.0%)	24 (57.1%)
#12b	非直接性	18 (42.9%)	3 (7.1%)	21 (50.0%)
#12c	不精確	11 (28.9%)	3 (7.9%)	14 (36.8%)
#12d	非一貫性	18 (42.9%)	2 (4.8%)	20 (47.6%)
#13	要約尺度	8 (23.5%)	6 (17.6%)	14 (41.2%)
#14a	研究結果の統合方法の記述	- -	- -	- -
#14b	一致性の尺度の記述	- -	- -	- -
#15a	臨床試験登録の検索	2 (4.8%)	30 (71.4%)	32 (76.2%)
#15b	著者への問合せ	1 (2.4%)	39 (92.9%)	40 (95.2%)
#15c	(事後メタアナリシス時)ファンネルプロット	- -	- -	- -
#15d	研究内での選択的報告及びその他の記述	13 (34.2%)	8 (21.1%)	21 (55.3%)
#16	追加的解析	- -	- -	- -
#17	研究の選択	0 (0.0%)	0 (0.0%)	0 (0.0%)
#18	研究の特性	7 (16.7%)	0 (0.0%)	7 (16.7%)
#19	研究内のバイアス・リスク	11 (26.2%)	1 (2.4%)	12 (28.6%)
#20a	各介入群の単純な要約データの記述	18 (45.0%)	5 (12.5%)	23 (57.5%)
#20b	フォレストプロット	- -	- -	- -
#21	結果の統合	- -	- -	- -
#22	全研究のバイアス・リスク	14 (33.3%)	0 (0.0%)	14 (33.3%)
#23	追加的解析	- -	- -	- -
#24	エビデンスの要約	25 (64.1%)	7 (17.9%)	32 (82.1%)
#25a	研究レベルとアウトカムレベルでの限界の記述	16 (38.1%)	13 (31.0%)	29 (69.0%)
#25b	レビューレベルでの限界の記述	17 (40.5%)	8 (19.0%)	25 (59.5%)
#26	結論	26 (61.9%)	6 (14.3%)	32 (76.2%)
#27a	SR の資金源と、その他の支援	8 (19.0%)	5 (11.9%)	13 (31.0%)
#27b	SR における資金提供者の役割	3 (7.5%)	8 (20.0%)	11 (27.5%)

※不備数(率) : 「不履行・記述なし」と「不十分又は不明瞭な記述」の合計数(率)。不備率は届出 SR 計 51 編から非該当を除いたものを分母として算出。

第4章 「機能性表示食品」制度における届出 SR に特有の報告方法の質の検証

本章では、「検索」、「個々の論文のバイアス・リスク評価」、「エビデンス総体の評価」、「メタアナリシスの手順・記述」、「その他の本制度の届出 SR に特有な内容」の適正性を、それぞれ届出 SR の記述内容から検証した。

各項目の評価結果と今後の課題について、以下に記述した。

第1項 検索に関する適正性の検証

図1の手順で抽出された51編を対象に、検索に関する適正性の検証を実施した。検証作業中に修正・更新が行われた場合は、修正・更新後の届出 SR に対して検証を実施した。

「検索に関する基本項目の記述」の検証に当たり、「検索に関する適正性検証リスト：機能性表示食品の届出 SR 用」を作成し、各項目について該当する届出 SR の編数と率を算出した。また「検索式（検索戦略）」については「検索式の評価事項」を作成し、評価を行った。

検証結果と今後の課題は以下のとおりである。

1. 検索に関する基本項目の記述の検証

1) 検索者

検証した届出 SR の過半数で、検索者の検索能力・経験が記述されていなかった。ガイドラインでは、サーチャー（情報検索技術者）や司書資格保持者など文献検索専門家による検索を義務付けてはいない。しかし、数多くのデータベースを駆使した網羅的な検索は、届出 SR の質を高める上で重要で、かつ高い専門性を要する作業であるため、専門家の参画が推奨される。

2) 検索日

検索日を正しく記述することが重要だが、一部の届出 SR において、検索日時が著しく古い場合があった。SR を行うに当たっては、最新の情報を収集することも重要である。

3) 検索情報源

届出 SR のほとんどが PubMed を利用していた他、JDreamIII、医中誌 Web など多様なデータベースが利用されていた。ガイドラインにおいては、検索するデータベース等の種類は特に定めていないが、「当該分野の文献検索で客観的に妥当と思われるものを適切に選定する」とこととしている。また、分野によっては PubMed に記載されていない論文の存在を指摘している。実際、SR などの網羅的なエビデンス収集の際には、可能な限り複数のデータベースや臨床研究登録システムを検索することが重要である。

4) 検索式

ガイドラインで準拠すべきとされている PRISMA 声明チェックリスト（2009年）においては「検索を再現できるよう、少なくとも一つのデータベースについての電子的な検索式 [search strategy] を、用いた全ての制限も含めて詳細に記述する」と記述されている。評価対象とした届出 SR 51編のうち49編で検索式が明記されていた。残りの2編については、検索式の一部が記述されていないといった不備があった。しかし、記述のあった届出 SR の中には、式の記述が不正確なものや、記述された検索式で検索をしても検索結果件数が一致しないものもあった。検索式の正しい記述はもちろん、検索結果編数を正確に記述することも重要である。

2. 検索式（検索戦略）の検証

検索式（検索戦略）の適正性について評価したところ、評価は届出 SR によってばらつきが見られた。これは、検索者の文献検索能力の違いが大きく反映されていることによると考えられる。キーワードやシソーラス^{注2}を適切に組み合わせ、データベースの特性に合わせて最適な検索式を設定している届出 SR もあったが、検索キーワードが不足している届出 SR や、検索対象年やアウトカム等で不必要に絞込みを行い、検索式を正しく記述していない届出 SR もあった。

ガイドラインでは、「網羅的な検索が実施できるよう、自由語及び統制語^{注3}（PubMed における MeSH^{注4}等）を適切に組み合わせた検索式を、文献データベースごとに設定する」、「言語バイアスを避ける観点から、少なくとも英語論文と日本語論文の検索を行う」とされている。SR の検索は、CQ（Clinical Question）ごとにキーワード、シソーラス（MeSH など）を組み合わせた最適な検索式をデータベースの特性に合わせて設定することが重要である。また、SR や臨床研究を検索する場合、PI(E)CO^{注5}の P、I(E)、(C)を用いて網羅的に行うが、害と不利益を含んだ幅広いアウトカムを拾うため、検索式に O は含めないことが望ましい。さらに、論文数が多い場合でも、言語などのフィルターを用いての絞込みは原則として行わないことが望ましい。ただし、プロトコールにあらかじめ記述されている場合には、作業量に応じて絞り込みを行うことも想定される。

第2項 個々の論文のバイアス・リスク評価の適正性の検証

届出 SR に使用された個々の論文のバイアス・リスク評価の適正性を検証した結果、届出 SR の中には、バイアス・リスクの評価方法について十分な記述がなされていないものがあることが分かった。評価方法に関する十分な記述がなければ、第三者が届出 SR を再現できないことに加え、得られた結果を客観的に判断することもできない。ガイドラインには明記されていないが、バイアス・リスクの評価が独立した 2 名で行われたか否かは、より適切に対象論文を精査したか否かを示す事項でもある。なお、本来ならば評価者 2 名の評価の一致度又は κ 係数^{注6}を記述する必要があるが、検証対象の届出 SR の中で、これらの記述があるものはごく一部であった。さらに、約 90% の届出 SR において、バイアス・リスクの評価結果に関する何らかの記述がなされているものの、「別紙の通り記述した」という不十分な記述に留まるものが存在した。

バイアス・リスクの評価は、方法及び結果に詳細を記述するとともに、考察においても十分に触れられるべき項目であるが、明確に説明されていないものが存在した。

個々の論文のバイアス・リスクの評価については、その評価方法はもちろん、評価結果を届出 SR 対象論文の採用基準にどのように反映させたか（バイアス・リスクがどの程度のものまでを採用したか）も明記することが望まれる。

^{注2} 索引付けや情報検索のための統制語の用語集のこと。用語の定義づけがなされており、階層構造により用語の上位・下位関係や、類義関係などが示されている。

^{注3} 「自由語」は検索者の任意で決められる言葉（フリーワード）を指す。自由語による検索は、語形や表記の変化、同義語、類義語等が存在することによる検索漏れやノイズが生じる懸念があるため、それらをリスト化し、代表として設定される語が「統制語」である。文献検索システムによって異なる場合もある。

^{注4} Medical Subject Headings の頭文字で、米国国立医学図書館が作成する生命科学分野のシソーラスのことで、毎年改訂されている。MEDLINE に収録される文献には、その内容に応じて MeSH により索引付けが行われている。

^{注5} 参加者（Participants）、介入（Interventions）または曝露（Exposures）、比較（Comparators）、結果（Outcomes）の頭文字を取ったもの。PICO、PICOS などと表記される場合がある。

^{注6} 二人の評価者間の評価の一致度を示す指標。0～1.0 の値を取り、値が大きいほど一致度が高い。

第3項 エビデンス総体の評価の適正性の検証（非直接性、不精確、非一貫性）

エビデンス総体の非直接性、不精確、非一貫性のそれぞれについて、評価方法が十分に記述されていた届出 SR は 17.6%、21.6%、29.4%、評価結果が十分に記述されている届出 SR は 37.3%、17.6%、41.2%にとどまり、第三者がエビデンス総体の評価内容を十分に把握し、理解できる届出 SR が多いとは言い難い状況であった。さらに、バイアス・リスク、非直接性、不精確、非一貫性の全項目で、方法と結果ともに十分な記述がそろっている届出 SR はわずか 4 編（7.8%）であったことから、届出 SR の中には、結論を客観的に判断し得る十分な情報が記述されていない場合があることがうかがえる。

中には届出 SR の対象論文が 1 報のみであったにもかかわらず、不精確や非一貫性の評価結果を本文、又は別紙様式（V）-13a（以下「別紙様式」は省略。）に記述した例もあり、その評価方法について疑問が残った。このことから、非直接性、不精確、非一貫性の評価方法については、より一層透明性を確保することが望まれる。

さらに、非直接性、不精確、非一貫性への評価を、エビデンス総体の評価にどのように反映したのかを明確に記述した届出 SR は 30%未満であり、この点も含めて評価内容の詳細を本文中に記述することが望まれる。

第4項 メタアナリシスの手順・記述の適正性の検証

メタアナリシスの手順及び記述に関する充足割合はおおむね高い値であった。これは、分析実施者がメタアナリシス専用の統計ソフトを使用していたことによると考えられる。ただし、統計ソフトを使わずに手計算されたと思われる事例もみられた。メタアナリシスの手法については、同一の分析業務受託者が実施していたためか、変量効果モデル^{注7}（DerSimonian-Laird Method）が 9 編中 6 編を占めていた。

今後の課題として、平均値・標準偏差の情報が欠落しているような研究をどのように取り扱ったかについての記述が求められる。特に、メタアナリシスに用いた研究論文の中には、平均値は記述があるものの、標準偏差は欠落しているものもあると考えられる。そういった研究論文は存在しなかったのか、95%信頼区間から算出したのか、又は除外したのか、平均値で埋め合わせをしたのか等の対処法についても記述が望まれる。また、原則としてメタアナリシスには RCT のみを併合すべきであり、そうでないときには感度分析（RCTに限ったかどうかなど）を実施すべきであろう。今回は問題なかったが、一致性が認められなかった場合、出版バイアスが疑われた場合、そして感度分析で結果の安定性が疑われた場合などには、一層留意してメタアナリシスを実施することが望まれる。

また、メタアナリシスを実施する際には、表 4 に示した 8 項目について十分な記述が望まれる。

^{注7} 全ての研究における結果のばらつきは、偶然誤差と研究ごとのバイアスが原因であると仮定した数理モデル。

表 4 メタアナリシスの報告に含めたい 8 項目

- 1) 併合対象となる指標(平均差、オッズ比など)
- 2) 併合手法(変量効果モデルなど)と研究に付す重み
- 3) フォレストプロットの描写
- 4) 異質性の検討(Cochran の Q 統計量、I² 統計量など)
- 5) 出版バイアスの検討(ファンネルプロットなど)
- 6) 感度分析(サブグループ解析など)
- 7) 用いた統計ソフト
- 8) 業務委託先

第5項 その他の本制度の届出 SR に特有な内容の適正性の検証

1. レビューワーカーの特性

届出 SR の中には、専門司書や博士の学位取得者、EBM (Evidence-based Medicine、科学的根拠に基づく医療) の専門家の参画が見られないものも存在した。ガイドラインにおいては、司書や学位取得者、EBM 専門家の参画を義務付けてはいない。しかし、数多くのデータベースを駆使した網羅的な検索と検索した論文のレビュー作業に専門知識を有する者が参画することは、届出 SR の質を高める上で重要な役割を果たすと考えられるため、推奨される事項であろう。

ただし、専門家であるがゆえにその責任も大きく、また利益相反にも関連する事項であるため、参画した場合にはその者の氏名や所属を明確にすることが望まれる。

2. ハンドサーチの実施

ハンドサーチを実施したことが記述されていた届出 SR は 12 編 (23.5%) だった。ハンドサーチは、学術雑誌や学会抄録等を用いて簡便にできる方法であり、対象論文の見落としを防ぐためにも、必要な検索方法である。ハンドサーチを行った場合は、学術雑誌等の名称・巻数・号数・出版年の記述や、ハンドサーチの実施方法などの明確な記述が必要である。

3. スクリーニング者の独立性

過半数の届出 SR において、文献のスクリーニングが 2 名以上で独立して実施されていた。SR はそもそも複数で行う研究である。それは、「人はケアレス・ミスをするもの」という前提に立っているからであり、ダブルチェックの重要性を示すものである。採用文献のスクリーニング段階で適格基準に合致した文献が漏れてしまうと、以後はそれに気付くことなく、分析へと移行してしまい、誤った SR になる可能性が高い。したがって、スクリーニングにおいては、2 名以上で独立して行うことを必須にすべきである。

4. 採用した研究デザイン

14 編 (27.5%) の届出 SR で、方法に研究デザインの記述がない、又は不明瞭であった。方法において PI(E)COS 形式での記述は重要であり、特に本文中に採用する予定の研究デザインを記述することは必須である。そうでなければ、機能性を示す研究結果が得られた研究デザインが選

択的に採用される危険が生じるであろう。そのため、SRの事前登録とも関係するが、採用する研究デザインを明確に決めてから、文献検索を行う必要がある。

5. 疾病罹患患者データ採用の有無

届出SRの中には、表や本文中から特定保健用食品の試験方法として記述された範囲内としての軽症者という水準を逸脱した、疾病罹患者が対象者に含まれているとも解釈できる論文を含めていることが判明したものがあつた。また、対象者に疾病罹患者しか含まれていない事例も存在した。軽症者と疾病に理解していない者の両方が含まれているものの、疾病に罹患していない者のデータのみを対象としたSRが別途実施されていない、又は実施した結果が詳細に明記されていない届出SRも存在した。

機能的表示食品は疾病の治療や予防を目的とするものでなく、健康の維持及び増進に役立つ旨又は適する旨を表示するものであるため、疾病罹患者のデータを届出SRに用いることはできない。科学的前提として、参加者の主要アウトカムや副次アウトカムの初期値が著しく高い又は低いものを含んでいる場合には、疾病罹患者データが混在していると考え、その研究は届出SRに使用すべきではない。

ガイドラインでは一定の範囲内で軽症者を取り扱うことを容認しているが、その際には対象者に必ず疾病に罹患していない者を含めた上で「疾病に罹患していない者のデータのみを対象とした研究レビューも併せて実施し（中略）その結果を、研究レビュー報告書及び研究レビューに関する一般消費者向けの抄録の両方に報告すること」とされている。

そのため、最低でも本文中の結果と（V）-11、（V）-13において、それぞれを独立して記述し、さらに考察において総合的に議論をするべきである。また、参考文献に疾病罹患者が対象者に含まれる論文等を用いた場合には、届出SRに使用した論文と明確に区別しなければならない。

いずれにせよ、疾病罹患者かどうかの判断がしづらいデータを届出SRに使用すると議論が複雑になるため、可能な限り届出SRの研究計画に含めるべきではない。

第5章 機能性表示食品の届出 SR の改善のポイント

本章では、前章までの検証結果を踏まえて、機能性表示食品の届出 SR を適正に作成するために、記述のポイントと具体例を示すとともに、実施すべきではないこと（不適正・研究倫理に関する事項）や、注意を要する事項を取りまとめた。

第1項 適正な記述方法の指針と注意

届出者の立場になれば、「検証結果は分かったが、具体的にどのように書けばよいのか」ということが最大の関心事のはずである。そこで、「【付録】PRISMA 声明チェックリスト：機能性表示食品のための拡張版」に基づく適正な研究レビューの記述例において、適正な記述方法の例を示した。定性的・定量的を問わず、ほとんど全ての届出 SR をカバーできるように、例示も「メタアナリシスを実施していない SR の場合」、「メタアナリシスを実施した SR の場合」と区別して記述した。届出 SR 全体の質を高めることを目的に、記述例として示すものであり、届出資料作成時の参考としていただきたい。

本事業の評価対象となった届出 SR の中には、PRISMA 声明チェックリストにおける項目番号と項目名（#13 要約尺度、などのように）を本文中に示しているものがいくつかあった。通常の学術論文としての SR では、投稿時に添付する別紙の PRISMA 声明チェックリストに記載ページを記述するスタイルが一般的であるが、本文中に項目番号と項目名が表示されていると、報告のポイントを明確に判断しやすいため、本制度における届出 SR においては推奨される記述方法である。そのため、届出を検討している届出者は、PRISMA 声明チェックリストの項目番号と項目名を本文中に入れることを推奨したい。そのことにより、確実に報告の質が高まるといえるだろう。

また適正な届出のためには、臨床試験でも SR でも、研究計画を適正に立てることが重要であり、これにより研究の質がほぼ決定付けられるといえる。

第2項 不適正・研究倫理に反すると考えられる注意事項

意図的でない場合でも、不適正・研究倫理に反しないために、絶対に行ってはならない事項や、逆に必ず行うべき事項がある。届出 SR において生じうる問題事項とその対処の具体例については、報告書本文を参照していただきたい。

第3項 その他の注意事項

その他の注意事項として、特に表示しようとする機能性の内容を記述する際の留意点などについてまとめた。

機能性の表示で最も重要なことは、PI(E)CO に基づいて得られた知見を正確に反映することである。その考え方の基礎は、P（参加者）においては「性・年齢・身体特性など、レビューした一次研究からどういった対象について言えるのか（例えば女性だけ、若年者だけなど、限定的ではないのかを適正に判断することは必須）」、I（介入）・E（曝露）においては「どの程度の摂取量で有効性がみられるのか、効果の発現にはどのくらいの時間が掛かるのか、用量反応関係はどうか」、C（比較）においては「比較するのに適切な食品・成分か」、O（アウトカム）においては「評価項目と表示内容は一致するのか、その評価項目からその文言を示してよいのか、拡大解釈しすぎていないか、多重仮説などにより偶然に得られた有意な結果ではないか、どこの身体部位での評価

なのか（ある部位での評価結果なのに、全身での評価と誤認されないか）」などである。

これらは基本事項であるが、大前提として届出 **SR** 自体が後付けではなく事前に綿密に立てられた研究計画の下で作成される必要がある。そして、作成された届出 **SR** は、不都合な研究結果も含めて漏れなく一次研究のデータを収集する一方、バイアスの高い一次研究を取り除く等、強固なエビデンスになり得る一定水準以上の一次研究結果から導き出されたものでなければならないことを忘れてはならない。

以上を踏まえて、届出者は消費者に普遍性のある正しい情報を伝えるとともに、消費者に誤認が生じないような表示内容にしなければならない。

第6章 おわりに

本報告書では、届出のあった食品の届出 SR の現状分析から課題点を見付け出し、今後、より適正な SR が実施されるための提案を行うことができた。

津谷（2003）^{注8}が示すように、「エビデンスには、つくる・つたえる・つかう、の3つの局面」がある。本制度に当てはめると、「つくる」に該当するものは一次研究、「つたえる」に該当するものは届出 SR、「つかう」に該当するものは食品に関する行政担当者、食品企業、一般消費者等ということになる。届出 SR の意義は、「一次研究の正しい結果を正しく伝えること」にある。届出者は本報告書を参考にして適切に計画・実施・記述することで、最低限の情報を含み、バイアスの少ない一定の質を有する SR を行うことができると考える。学術研究における SR のようなハイレベルを極めるということでなく、食品であるがゆえの難しさもしんしゃくしたベーシック・スタンダードレベルということで、本報告書には付録として「PRISMA 声明チェックリスト：機能性表示食品のための拡張版」に基づく適正な研究レビューの記述例を添付し、具体的な提案や例示をしている。SR の総論・各論に不安や疑問を抱いている届出者にとって役立つであろう。

末筆になるが、本制度が開始されてようやく1年を迎えたということは、言い換えれば齢にしてまだ満1歳ということである。まだ発展過程にあり、関係各位が知恵を出し合い建設的な議論を積み重ね、研究者・研究機関が教育的な支援を続け、届出者（SR 委託業者も含む）が高い倫理観（企業倫理・研究倫理）を持って誠実で丁寧な消費者への情報提供に取り組み、本制度の充実を図っていくことが重要である。

^{注8} 津谷喜一郎.EBM におけるエビデンスの吟味. Therapeutic Research, 2003;24:1415-1422.