

# 機能性表示食品制度における システマティック・レビューの質の再検証 —PRISMA 2020 の導入を目前にして—

上岡洋晴<sup>1</sup> 津谷喜一郎<sup>2</sup> 折笠秀樹<sup>3</sup>

## はじめに

2015年4月に開始された機能性表示食品制度は2023年4月をもって9年目に入る。2022年12月31日時点において、消費者庁のホームページにおける届出数は6,213件と膨大な数になっている。機能性のエビデンスとしては5,872件(94.5%)が研究レビュー、そして残りが臨床試験(clinical trial: CT)の結果を用いている。この研究レビューだが、「機能性表示食品の届出等に関するガイドライン」<sup>1)</sup>において、2009年に公表された初版のPRISMAチェックリスト<sup>2,3)</sup>に準拠するように求めていることから、事実上、システマティック・レビュー(systematic review: SR)を意味している。以後はSRと称して論を進める。

本制度は企業等の届出者の責任に基づく届出制度であることから、当初からエビデンスの資料の質に関して消費者団体やマスコミなどから疑問を呈する声が挙げられていた。アカデミア研究者においても類する見方があった。機能性表示食品は英文で“Foods with Function Claims (FFC)”という表記なので、著者らは「届出SRをさらにSRする」ということで、2015年3月にFFC-SR2という名の研究グループ(ステアリング・オーサー: 上岡洋晴, 津谷喜一郎, 折笠秀樹)を立ち上げた。

さて、個々の機能性を示すCTとSRについては「研究の質」(研究方法論として質のこと)と「報告

の質」(記載内容の質)ともに改善すべきことがアカデミア研究者から複数の英文誌等で報告されている。表1は本制度におけるエビデンスの質を高めるための主な調査・研究報告の一覧である。消費者庁による検証事業は2回(SRとCTについて1回ずつ)、アカデミア研究もSRとCTともに異なる観点から着手されてきた。これらの詳細は後述する。

CTに関しては、CONSORT 2010チェックリスト<sup>4,5)</sup>に基づいて十分な記載がされていないという2018年のTanemuraらの報告<sup>6)</sup>、プロトコールと論文中の内容が一致していないという2021年のKamiokaらの報告<sup>7)</sup>、バイアスリスクが高いという2022年のKamiokaらの報告<sup>8)</sup>がある。これらから改善ポイントを端的に要約すると、「研究計画のプロトコールに関して透明性をもって正確に記載するとともに、その段階でバイアスのない方法を適切にセットして実施すること、そしてチェックリストに基づいて丁寧に記載することが必須である。」と示すことができるだろう。

一方、SRに関しては、2007年に公表された初版のAMSTARチェックリスト<sup>9,10)</sup>に基づいて研究の質を評価した結果、かなり低いとする2017年と2019年のKamiokaらの報告<sup>11,12)</sup>、初版のPRISMAチェックリストに基づいて報告の質を評価した結果、不備が多いという上岡ら2019年<sup>13)</sup>の報告がある。こちらにも改善ポイントをまとめると「事前にPRISMAチェックリストに基づき、またAMSTAR

<sup>1</sup> 東京農業大学大学院環境共生学専攻 <sup>2</sup> 生存科学研究所 <sup>3</sup> 統計数理研究所

表 1 機能性表示食品制度における機能性の科学的根拠の質を高めるための主な調査・研究報告

(臨床試験：CT, システマティック・レビュー：SR と略す)

公表年月	消費者庁	アカデミア研究	その他重要な関連事項
2015年4月	機能性表示食品制度の開始		
2016年7月	【SR】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報告書 <sup>14)</sup>		
2016年12月		【SR】機能性表示食品のシステマティック・レビューのための教育図書発刊 <sup>15)</sup>	
2017年3月	【CTと安全性】機能性表示食品制度における臨床試験及び安全性の評価内容の実態把握の検証・調査事業報告書 <sup>27)</sup>		
2017年4月		【SR】AMSTARチェックリストに基づいてのSRの研究の質評価研究 <sup>11)</sup>	
2018年3月		【CT】CONSORT 2010チェックリストに基づいての報告の質評価研究 <sup>6)</sup>	
2018年11月			【SR】AMSTAR 2の日本語訳と解説 <sup>23)</sup>
2019年3月		【SR】PRISMAチェックリストに基づいての消費者庁の検証事業前後の報告の質の比較研究 <sup>13)</sup>	
2019年6月		【SR】AMSTARに基づいての消費者庁の検証事業前後の研究の質の比較研究 <sup>12)</sup>	
2019年12月		【SR】「PRISMA声明チェックリスト：機能性表示食品のための拡張版」における追加項目の開発 -PRISMA FFC extension ver.2- <sup>17)</sup>	
2021年6月			【SR】PRISMA 2020声明の日本語訳と解説 <sup>19)</sup>
2021年7月			【SR】PRISMA-Sの日本語訳と解説 <sup>21)</sup>
2021年12月		【CT】プロトコールと論文記載内容の一致度(コンプライアンス)を評価した研究 <sup>7)</sup>	
2022年9月		【CT】バイアスリスク評価研究 <sup>8)</sup>	
2022年11月		【SR】AMSTAR2による質評価のUMIN登録(詳細プロトコール入手先： <a href="https://1drv.ms/b/s!AtViZRADbyukiG7zFy8oShYcWxRF?e=xgKqjL">https://1drv.ms/b/s!AtViZRADbyukiG7zFy8oShYcWxRF?e=xgKqjL</a> )	
2023年1月		【SR】本論	
2023年3月	【SR】PRISMA 2020チェックリストをガイドラインへ反映(見込み) → 一定の猶予期間		
2023年X月	【SR】PRISMA 2020チェックリストに基づくSRでの届出開始(見込み)		

で研究の質にダメージを与える事項を回避するように研究計画を立てて実施・報告することが必須である」といえるだろう。しかし、残念ながら現状では改善が進まず、届出されるSRの数だけは増え続け

ている。

そこで本論では、制度開始後8年が経過しようとしている中でSRに焦点を絞り、これまでの研究と報告の質を高めるための各種調査や研究報告の経過

をナラティブ・レビューするとともに、今後の課題と展望を示すこととした。

### 制度開始初年度 (2015 年度) における SR の報告の質評価

2016 年 7 月 7 日に消費者庁から「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報告書」<sup>14)</sup> が公表され、SR の報告方法を中心として、その問題点や改善すべき事項が明らかにされた。これは、検証事業のワーキンググループが本制度開始直後の 2015 年 4 月 1 日から 10 月 31 日までの 7 ヶ月間に登録されたすべての SR (51 編) を対象として、PRISMA チェックリスト (2007 年版) をさらに再分割・本制度の SR に適応させた「PRISMA チェックリスト：機能性表示食品のための拡張版 (PRISMA 機能性表示食品のための拡張版)」に基づき、記述の有無や不備のある点をチェック項目ごとに示したものであった (表 2：左側部分)。

ここでは PRISMA チェックリスト 45 項目それぞれについて「不履行・記述なし」と「不十分または不明瞭な記載」を「不備」とした。その項目ごとの不備率のラベリングとして「A：0 から 20% 未満の SR で不備がある」「B：20% 以上 40% 未満の SR で不備がある」「C：40% 以上 60% 未満の SR で不備がある」「D：60% 以上の SR で不備がある」とした。表 2 をみると総じて不備の多い SR が多いことがわかる。とくに D と評価されたのが「構造化抄録の書き方」「プロトコルの事前登録」「臨床試験登録の探索」「著者への問い合わせ」「エビデンスの要約」「研究レベルとアウトカムレベルの限界の記述」であった。

不備に関しては、評価者にとって判定が難しい書きぶり、たとえば、ある項目に関して「記載はあるといえはあるが、十分な情報とはいえない」とか「記載はあるが、的外れな内容になっている」というような SR が少なくないことが明らかにされている。

SR はある介入 (機能性関与成分) の機能性を示す集大成であり、不明瞭な記述は避け、必要な情報がすぐに把握できる SR でなければならないことが報告書の中で述べられている。さらに報告書には付

録として、その後の良好な SR づくりのために「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が 17 ページにわたって示されている。これを下敷きにして作成すれば確実に前述のような不備は少なくなり一定水準以上の SR 報告ができるものと期待された。

### 制度開始 3 年目 (2017-2018 年度) における SR の報告の質

前述の報告書<sup>14)</sup>に加え、2016 年 12 月には本制度の SR 実施のための教育図書<sup>15)</sup>が発刊され、届出において好ましくない記載例や適正な例を織り交ぜた内容などが提示された。本制度に興味や関心のある企業をはじめ多様な業界団体、SR の代行実施業者、アカデミア研究者などに食品としての適切な SR 報告の方法が周知されたと考えられた。制度開始から 2 年以上が経過し、新規の届出 SR は当初の頃の SR よりも報告の質は改善・優良なものになっているだろうと推測された。

そこで、FFC-SR2 は、2018 年に検証事業実施後に SR の報告の質が向上したかどうかを「PRISMA チェックリスト：機能性表示食品のための拡張版」を用いて明らかにすることを目的として検証事業の実施前後の比較研究を行った<sup>13)</sup>。消費者庁ウェブサイトに掲載された機能性表示食品の届出 SR を対象とし、検証事業と同じく期間 7 ヶ月として 2017 年 7 月 1 日から 2018 年 1 月 31 日までに掲載されたすべての SR104 編 (重複した SR は除外) を対象とし、「検証事業後 SR」とした。比較として検証事業で用いた 51 編を「検証事業前 SR」とした (図 1)。

「検証事業後 SR」としての期間設定の理由として、検証事業の報告書は 2016 年 7 月 7 日に公開されたが、前述の組入れ開始となる 2017 年 7 月はそれから 1 年経過を意味しており、届出者等に浸透するのに十分な期間だと考えられた。つまり、届出者が作成中あるいは計画中の SR に検証事業における指摘事項を反映させて良好な報告にするために十分な猶予期間であると判断した。

検証事業前後を含めて、表 2 に全体の結果を示した。表の右側が検証事業後 SR である。45 項目の不

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」に基づく記述の不備 (文献 13,14 より作表)

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	検証事業前 届出SR (n=51)			検証事業後 届出SR (n=104)			P value
		不備			不備			
		度数	率	ラベリング**	度数	率	ラベリング**	
#1	タイトル	0	0	A	1	1.0	A	1.00
#2	構造化抄録	33	64.7	D	98	94.2	D	<0.01
#3	論拠	25	49.0	C	73	70.2	D	0.01
#4a	参加者の記述	6	11.8	A	29	27.9	B	0.02
#4b	介入の記述	6	11.8	A	19	18.3	A	0.36
#4c	比較の記述	8	15.7	A	21	20.2	B	0.66
#4d	アウトカムの記述	6	11.8	A	33	31.7	B	<0.01
#5a	レビュー・プロトコルの有無	26	51.0	C	40	38.5	B	0.14
#5b	プロトコルへのアクセスの可否	49	96.1	D	100	96.2	D	1.00
#5c	プロトコルのweb address, 登録番号の有無	49	96.1	D	102	98.1	D	0.60
#6a	研究の特性の記述	17	33.3	B	56	53.8	C	0.02
#6b	報告の特性の記述	22	43.1	C	46	44.2	C	0.90
#7	情報源	22	43.1	C	60	57.7	C	0.09
#8	検索	8	15.7	A	13	12.5	A	0.62
#9a	スクリーニング方法に関する記述	21	41.2	C	43	41.3	C	0.98
#9b	適格性に関する記述	12	23.5	B	29	27.9	B	0.56
#9c	採択基準に関する記述	17	33.3	B	29	27.9	B	0.49
#10	データ収集プロセス	24	47.1	C	61	58.7	C	0.17
#11	データ項目	13	25.5	B	18	17.3	A	0.23
#12a	バイアスリスク	29	56.9	C	83	79.8	D	<0.01
#12b	非直接性	26	51.0	C	83	79.8	D	<0.01
#12c	不精確	16	31.4	B	84	80.8	D	<0.01
#12d	非一貫性	22	43.1	C	85	81.7	D	<0.01
#13	要約尺度	16	31.4	B	30	28.8	B	0.75
#14a	研究結果の統合方法の記述*	4	44.4	C	0	0.0	A	0.02
#14b	一致性の尺度の記述*	2	22.2	B	1	7.7	A	0.54
#15a	臨床試験登録の検索	40	78.4	D	72	69.2	D	0.23
#15b	著者への問合せ	47	92.2	D	92	88.5	D	0.58
#15c	(事後メタアナリシス時)ファンネルプロット	1	11.1	A	1	7.7	A	0.55
#15d	研究内での選択的報告及びその他の記述	24	47.1	C	86	82.7	D	<0.01
#16	追加的解析*	3	33.3	B	2	15.4	A	0.61
#17	研究の選択	1	2.0	A	5	4.8	A	0.66
#18	研究の特性	10	19.6	A	39	37.5	B	0.02
#19	研究内のバイアス・リスク	14	27.5	B	15	14.4	A	0.08
#20a	各介入群の単純な要約データの記述	24	47.1	C	26	25.0	B	<0.01
#20b	効果の推定量と信頼区間の記述*	2	22.2	B	1	7.7	A	0.54
#21	結果の統合*	1	11.1	A	1	7.7	A	1.00
#22	全研究のバイアス・リスク	17	33.3	B	48	46.2	C	0.13
#23	追加的解析*	5	55.6	C	1	7.7	A	0.02
#24	エビデンスの要約	39	76.5	D	89	85.6	D	0.18

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」に基づく記述の不備 (続き) (文献 13,14 より作表)

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	検証事業前 届出SR (n=51)			検証事業後 届出SR (n=104)			P value
		不備			不備			
		度数	率	ラベリング**	度数	率	ラベリング**	
#25a	研究レベルとアウトカムレベルでの限界の記述	33	64.7	D	70	67.3	D	0.75
#25b	レビューレベルでの限界の記述	30	58.8	C	49	47.1	C	0.17
#26	結論	39	76.5	D	86	82.7	D	0.39
#27a	SRの資金源と、その他の支援	15	29.4	B	54	51.9	C	<0.01
#27b	SRにおける資金提供者の役割	12	23.5	B	55	52.9	C	<0.01

\*メタ分析を実施したSRのみが対象

\*\*不備率として、A:0から20%未満、B:20から40%未満、C:40から60%未満、D:60%以上のSRで該当

$\chi^2$ 検定による有意確率。セルが10未満を含む場合にはフィッシャーの直接確率計算法により算出

備率のラベリングは検証事業の方法と統一している。有意確率は前後比較の検定であり、有意に向上したか悪化したかの判断指標となる。なお、メタアナリシスのありとなしを分割した結果と個々のSRの評価結果は紙面の都合のため割愛した(分割した解析結果は文献 13にあるのでそちらを参照のこと)。

表 2 の見方は、たとえば「#1 タイトル」において「SR」か「メタアナリシス」か、あるいは両方であるかが明確にわかるように記載することが求められているが、前後ともに不備率は低く (A)、この点においては良好であることを意味している。「#2 構造化抄録」においては、背景・目的・検索で用いたデータ源・研究の適格基準・PICO・研究の評価と統合法・結果・限界・結論重要な知見の意味合い・SRの登録番号を含めることが求められているが、ラベリングはDであり、検証事業後の方が不備率が有意に増加していた。

このように個々のチェック項目をみていくと、依然としてラベリングがCやDのものが多く、むしろ不備率が高まっているという仮説と相反する結果となった。すなわち、報告の質の低いSRが多く、消費者庁の検証事業報告書や教育図書によって改善がなされているとはいえないという大変衝撃的な結果となっていた。

この理由を把握することが課題解決に向けてもつとも重要である。検証事業の報告書や前述の図書をよく読んでいない、あるいは理解できていないままでエビデンス資料として形式だけSRらしく設えて

提出してしまっていることが最大の原因だと考えられる。

さらに推測の域を出ないが、本制度の発足当初は食品業界の大企業の届出が中心で、ある一定レベルのCTに関する知見があり、まずまずのSRであったが、以後2016年から2018年にかけて爆発的な数の届出が行われ、質の低いSRを模倣したり、使い回したSRが公開されていた可能性があった。もっと具体的に述べれば、「ある機能性関与成分Xに関して、企業AのSRがすでに消費者庁ウェブサイトに掲載されているのだから、そのXを最終製品に入れて販売するには、そのSRをなぞってやればシンプルで問題なく受理されるだろう」ということで、別の複数の企業(B社、C社、D社……)がXについて低い水準のSRのままで出していることが質低下の理由の1つではないかと考える。

多くの届出者が共通して扱っている機能性関与成分においては、後発のSR届出者が、先行研究と位置づけられるSRを確実に見て、参考にしていると考えられる。したがって、それは参考文献あるいは引用文献に明確に示さないと盗用ともいえるものである。検索戦略なども然りである。サイエンスとして不適切であり、論文の基本中の基本たる作法は当然、踏襲すべきである。競合他社のSRであってもビジネス倫理にも反している。

また、利益相反 (conflict of interest: COI) の記載も不十分な届出がまだまだ少なくない。COIに関する事項がないなら「なし」と記載しなければならない。また、SRの資金源の記載やCROの関与、あるい

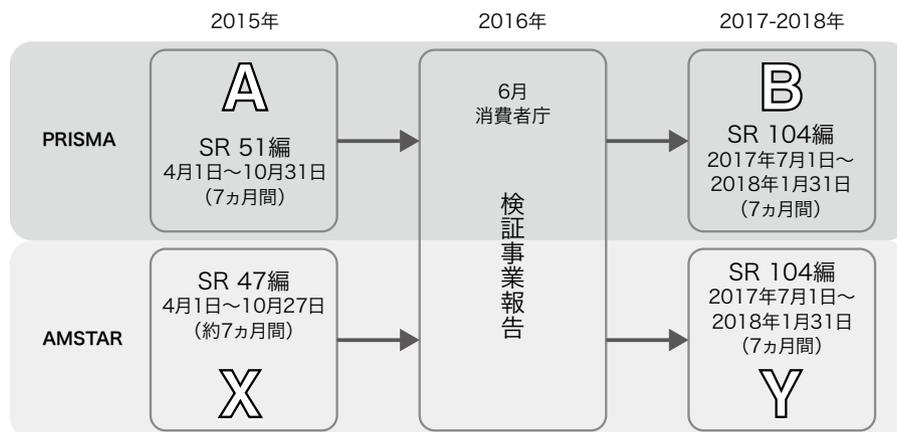


図1 消費者庁における検証事業前後比較の概要  
PRISMA：AとBの比較，AMSTAR：XとYの比較

はSR代行業者の関与，アカデミア研究者の所属名・氏名など，関係するすべての事項を明確にしなければならない。

本制度は届出制で内容に関する審査がないことも，企業努力をしない方向性を後押ししていると考えられる。津谷は「エビデンスの総体を対象とするSRには賞味期限がある。健康食品の承認・届出ともに再審査/再評価に類する制度があってもよいのではないか。」<sup>16)</sup> という見解を述べている。別の観点からのアプローチとしては，アカデミア研究者らの査読 (peer review) を受けること，すなわち論文文化して学術誌にアクセプトされることである。これにより，致命的な問題はなく，研究・報告の質が一定水準以上であることのお墨付きを得ることができる。

### 消費者庁の検証事業前後におけるSRの研究の質

SRの研究の質について，図1に示したような検証事業前後における比較研究のデザインでの報告<sup>11, 12)</sup> がなされた。「SRの研究自体の質」を評価するツールに関しては，前述したAMSTARチェックリストを用いた。「#1 デザインは“あらかじめ”提示されたか?」「#2 研究の選択とデータの抽出は複数で行われたか?」「#3 包括的な文献検索がなされたか?」「#4 出版形態 (灰色文献など) 組入れ基準に用いられたか?」「#5 (組入れられたおよび除外された) 研究リストは提示されたか?」「#6 組入

られた研究の特性は提示されたか?」「#7 組入れられた研究の科学的な質は評価され，記録されていたか?」「#8 組入れられた研究の科学的な質は，結論を導く際に適正に利用されたか?」「#9 研究結果を併合するのに用いられた方法は適正だったか?」「#10 出版バイアスの可能性は評価されたか?」「#11 利益相反は明示されたか?」の11項目で構成され，良好であれば各項目1点ずつで合計11点満点にて評価する。

前述のPRISMAチェックリストは，報告の質を高めるために開発されており，各項目における記述漏れがないようにすることが目的で，そもそも合計点での高低を評価するツールではない。一方，AMSTARチェックリストはもともと個々の項目で評価するのに加えて合計点であるSR全体の研究の質を評価することをねらいとして開発されていたので，そのとおり実施した。

研究の結果，合計得点において検証事業前SRでは6.2±1.8 (平均±標準偏差) 点であり，検証事業後SRでは5.0±1.9点で研究の質が有意に低下していた。PRISMAチェックリストとまったく同様な傾向を示していた (個々の項目の結果は割愛)。これらは，「研究の質も，報告の質も，改善しておらず，質の低いSRが多く届出されている」ことを示している。2018年から現在にかけてのSRは評価していないものの，同様な傾向にあるのではないかと推測する。

## PRISMA 2020 で届出 SR の質は向上するのか？

時系列でみると PRISMA チェックリストに対して本制度特有のチェック項目を追加した「PRISMA 声明チェックリスト：機能性表示食品のための拡張版」(PRISMA FFC extension ver.2)<sup>17)</sup> が 2019 年 12 月に公表された。そこにはガイドラインでの特異的な要求事項の併記に加え、「他の食事や身体活動・運動との併用を行っての介入の場合にはその全体像を示すことや、その相互作用の影響を考察すること」「機能性関与成分の消化吸収に関する根拠を示すこと」「生鮮食品の SR の場合には AGRIS (Agricultural Science and Technology Information) を文献検索データベースとして加えること」などが重要項目として追加され、その記載が推奨されている。

さて、2009 年に初版の PRISMA 声明が出て以来、10 年以上が経過した。人工知能 (artificial intelligence: AI) の普及は急速で、この間、オープンアクセス誌の増加やヘルスケアのあらゆる分野における SR の数の急増が生じ、その質が問われている中で待望の PRISMA 2020 声明<sup>18,19)</sup>が公表された。初版の PRISMA (2009) は完全に刷新されているので (本文中にも初版 PRISMA は使用するべきではないとの記載)、今はこの 2020 年版に従う必要がある。

PRISMA 2020 の特徴を一言でいえば、「透明性のために、すべて記載せよ」となる。そのポイントを示すと、「構造化抄録が新しくなったこと」「フローチャートが新しくなり、アップデート版の SR の場合には前の SR での検索との区別した記載が求められるようになったこと」「検索において少なくとも 1 つのデータベースではなく、すべてのデータベース登録サイトなどにおける検索戦略を明記すること」「スクリーニングで自動化ツールを用いた場合にはその詳細を示すこと」「メタアナリシスにおける実施方法と結果記載の明確化」「エビデンス総体の確実性 (certainty)・信頼性 (reliability) の評価方法と結果を明記すること」「利益相反を宣言すること」である。

とくにエビデンス総体 (totality of evidence) の確実性は、どれだけ厳格に当該 SR の信頼性を示せるかを問うものであり、「機能性ありき」の論理展開

をしてきた届出者においては厳しい選択 (すなわち効果推定量の信頼性は低いと自ら書けるのか) に迫られるかもしれない。

ヘルスケアにおけるアカデミア研究者は、さほど違和感なくこの 2020 年版を受け入れることができるだろうが、本制度においては懸念がある。前述のように初版の PRISMA でさえも理解できていない届出者が、新版に基づいて適正に記載できるかである。一方で、これまで優良な SR を作成してきた届出者においては苦慮なく乗り換えることができることも事実だろう。もはや初版は存在しないことから、近く改訂されるガイドラインにはこの新しい PRISMA 2020 チェックリストに基づいた SR を記載することが要件に盛り込まれるだろう。正しい理解のために有識者による潜在的な届出者のための研修会や相談会などが必要かもしれない。

なお、2021 年には SR の検索戦略に特化したチェックリストである PRISMA-S<sup>20,21)</sup> が公表された。IT を駆使した検索をどのように記載するかが詳細に示されている。また、検索で漏れが生じるとその必要な論文を再度入手するのは困難であり、この検索戦略は SR では極めて重要である。この PRISMA-S では、ある検索者の検索戦略を別の独立した検索の専門家が評価し、場合によっては協議して適正な検索式を導き出すピアレビューの導入を推奨している。周囲にこうした臨床疫学的研究に精通した検索者 (searcher) がいるかどうかは鍵となり、この実施は少しハードルが高いかもしれない。しかし、ヘルスケア研究における高品質の SR は、今後このような手続きを踏んで作成されるものであろう。PRISMA-S はその日本語訳版を参考にさせていただきたい。

ここで 1 つ強調したいことがある。本制度発足当初 2015 年に届出された SR は 8 年経過していることになる。「10 年ひと昔」から今は大きく変化した。ネット上では無料で全文にアクセスできる「オープンアクセス誌」が多くなり、掲載される CT の論文数も多くなっている。こうした状況下では、SR に新たな CT の結果を取り入れたアップデートが課題である。

表3 AMSTAR 2 における7つの重大領域と信頼性評価 (文献 22, 23 から抜粋)

7つの重大領域
#2 レビューの報告に、レビュー実施前にその方法を決定したことを明示し、重大な計画変更があれば正当な理由を示したか?
#4 レビュー著者は、網羅的な文献検索の手法を用いたか?
#7 レビュー著者は、除外した研究のリスト提示し、その除外が正当である理由を示したか?
#9 レビュー著者は、レビューに組み入れた各研究のバイアスリスク評価に際し、満足のいく手法を用いたか?
#11 メタアナリシスが行われた場合、レビュー著者は、結果の統計的併合において適正な方法を用いたか?
#13 レビュー著者は、レビュー結果の解釈や考察をする際にバイアスリスクを考慮したか?
#15 量的な統合がなされた場合、レビュー著者は出版バイアスについて十分な検討とレビュー結果に与える影響の可能性を考慮したか?
信頼性の具体的な4段階評価と解釈
「高(High): 重大な弱点はない、または1つだけ重大ではない弱点がある。」 →そのSRは関心のある疑問に対する利用可能な研究結果の正確かつ包括的な要約を提供している。
「中(Moderate): 2つ以上の重大ではない弱点があるが、重大な弱点はない。」 →そのSRは、組み入れられた利用可能な研究結果の正確かつ包括的な要約を提供しているかもしれない。
「低(Low): 重大ではない弱点の有無にかかわらず、1つだけ重大な弱点がある。」 →そのSRは、重大な欠陥があり、利用可能な研究結果の正確かつ包括的な要約を提供していないかもしれない。
「極めて低(Critically low): 重大ではない弱点の有無にかかわらず、2つ以上の重大な弱点がある。」 →そのSRは、重大な欠陥があり、利用可能な研究結果の正確かつ包括的な要約を提供していると信頼すべきではない。

### 最新のSRの質を問う 新たな研究プロジェクトの立ち上げ

FFC-SR2は、2022年11月19日に研究題目「日本の機能性表示食品において届出されたシステムティック・レビューについての横断研究: AMSTAR 2<sup>22, 23)</sup>による質評価」(代表研究者: 上岡洋晴)としてUMIN登録(UMIN000049558)を行った。詳細なプロトコルは次のURL (<https://1drv.ms/b/s!AtViZRADbyukiG7zFy8oShYcWxRF?e=xgKqjL>)から誰でも閲覧できる。

この研究は、機能性表示食品制度において機能性のエビデンスとして届出されたSRの研究の質についてAMSTAR 2を用いて評価し、SRの不適切な方法となっているその特徴を明らかにすることを目的としている。2022年4月1日から2022年10月31日までに消費者庁ウェブサイト公表された記号番号HシリーズのSRで、具体的にはH1-H535の届出で臨床試験の届出と撤回を取り除いた519編のうち、ランダム抽出した40編のSRを評価する予定である。

AMSTAR 2は16項目から構成されており、とくに次の7つの重大領域がある(表3)。また、4段階からなる全体の信頼性評価と解釈も示した。たとえば、プロトコルを登録していない時点で、重大領

域の#2が不適正ということになり、それだけで信頼性は「低」(Low)となる。さらにもし、レビュー結果の解釈や考察において、バイアスリスクに関して適切な説明がなされていない、つまりバイアスリスクの評価はしていても、その結果どのような手続きを行ったか、どのように解釈するのかの記載がないと、重大領域の#13が不適正となる。これで重大領域の2つが不適正となり、そのSRの信頼性は「極めて低い」(Critically low)ことになり「そのSRは、重大な欠陥があり、利用可能な研究結果の正確かつ包括的な要約を提供していると信頼すべきではない。」という結論づけがなされる。FFC-SR2としては、この厳格なSRの信頼性評価ツールを用いた場合、本制度におけるSRはどの程度のレベルであり、またどのような項目で不適正なのかを明らかにし、信頼の置ける届出にするための参考資料を提供したいと考えている。

### おわりに

本論ではSRをめぐる報告と研究の質について消費者庁の検証事業やアカデミア研究者の研究結果を中心に議論をしてきた。社会経済、とくに食品業界においては、本制度の利用により販売促進がなされ収益を高めることが成長戦略、規制緩和の実質的な

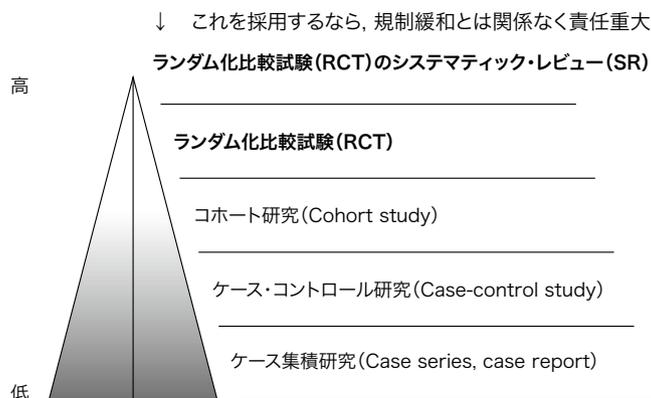


図2 エビデンス・グレーディングと本制度のSRの位置づけ

果実となるだろう。一方、消費者側からすれば、安全なことは当然ながら、機能性が人で本当に存在するのか、効果はあるのか、がもっとも大切である。

本制度はCTに加えてSRというエビデンス・グレーディングが最上位の研究デザインを採用している(図2)。結論はより真実を示す可能性が高いともいわれる研究デザインである。アカデミアによるSRも営利目的の組織団体企業等におけるSRも違いはなく、実施者はその機能性をバイアスなく適正に伝える義務がある。

本制度においては、SRに限ってはUMIN-CTRやPROSPEROなどへの事前登録を届出の要件とはしていない。CTは必須であるのに、エビデンス・グレーディングが最上位なSRがどちらでもよい、というのは極めて奇異である。本制度開始10年に差し掛かろうとしているが、届出者が率先して事前登録、詳細なプロトコルを公開して後付け解析などを行っていないことを示すときにきている。

サイエンスは進化し続けるものである。PRISMA 2020はよりよいSRの報告とバイアスを回避するのに有用なチェックリストであり、歓迎されるべきである。これを精読する必要があるが、PRISMA 2020解説と詳細<sup>24)</sup>とともに正しい理解が必要であることを申し添える。新規の申請や、すでに届出済みのSRのアップデートのために、今から入念な準備がなされることを期待したい。

なお、余談になるがCTに関連する最新情報として、CONSORTグループが中心となって、CONSORT 2010<sup>4,5)</sup>と、そのプロトコル拡張版で

あるSPIRT<sup>25,26)</sup>を更新するための研究が進められている。近い将来、これらもアップデートされ本制度にも反映されるものと考えられる。

本文中では引用しなかったが、表1で消費者庁の検証事業として「機能性表示食品制度における臨床試験及び安全性の評価内容の実態把握の検証・調査事業報告書」<sup>27)</sup>を紹介した。

**附記** 本論の資金源は、2022-2023年度日本学術振興会科学研究費基盤C(課題番号21K11604):研究題目「機能性表示食品における臨床試験とシステマティック・レビューの内的妥当性評価研究」代表研究者:上岡洋晴)であった。主宰者は上岡洋晴, 主宰団体はFFC-SR2である。

膨大な届出資料・データの収集・整理等を行っていただいた研究アシスタントの石田美千子氏に感謝の意を表する。

**利益相反** 著者上岡は、ファンケル(株)、カゴメ(株)、森下仁丹(株)、サントリー・ウエルネス(株)、丸善製薬(株)のSRの論文指導を行い、謝金を得た。他の著者においては、申告すべき事項はない。

## 参考文献

- 1) 消費者庁ホームページ. 機能性表示食品の届出等に関するガイドライン.  
[https://www.caa.go.jp/policies/policy/food\\_labeling/foods\\_with\\_function\\_claims/assets/foods\\_with\\_function\\_claims\\_220401\\_0002.pdf](https://www.caa.go.jp/policies/policy/food_labeling/foods_with_function_claims/assets/foods_with_function_claims_220401_0002.pdf)
- 2) Moher D, Liberati A, Tetzlaff J, Altman DG; PRISMA Group. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. PLoS Med 2009; 21; 6: e1000097.
- 3) 金子善博, 津谷喜一郎, 中山健夫. システマティック・レビューおよびメタアナリシスの報告における望ましい報告項目: PRISMA 声明. In: 中山健夫・津谷喜一郎編著. 臨床試験と

- 疫学研究のための国際ルール集 (Part 2). ライフサイエンス出版 ;2016.p.140-7.
- 4) Schulz KF, Altman DG, Moher D; CONSORT Group. CONSORT 2010 statement: updated guidelines for reporting parallel group randomised trials. *BMJ* 2010; 340: c332.
  - 5) 津谷喜一郎, 元雄良治, 中山健夫. CONSORT 2010 声明ランダム化並行群間比較試験報告のための最新版ガイドライン. *薬理と治療* 2010;939-49.
  - 6) Tanemura N, Hamadate N, Urushibara H. Evaluation of randomized controlled trials of Foods with Functional Claims request: The learning outcomes from studies in Japan. *J Funct Foods* 2018; 42: 248-53.
  - 7) Kamioka H, Origasa H, Kitayuguchi J, Tsutani K. Compliance of clinical trial protocols for Foods with Function Claims (FFC) in Japan: Consistency between clinical trial registrations and published reports. *Nutrients* 2021;14:81.
  - 8) Kamioka H, Origasa H, Kitayuguchi J, Yoshizaki T, Shimada M, Wada Y, et al. Risk of bias in clinical trials reported for Foods with Function Claims in Japan: a cross-sectional study on research quality. *J Clin Trials* 2022;12:1000503.
  - 9) Shea BJ, Grimshaw JM, Wells GA, Boers M, Andersson N, Hamel C, et al. Development of AMSTAR: a measurement tool to assess the methodological quality of systematic reviews. *BMC Med Research Methodol* 2007;15:7:1-7.
  - 10) 折笠秀樹, 上岡洋晴, 津谷喜一郎. AMSTAR: システマティック・レビューの方法論的な質を評価するための測定ツールの開発. In: 中山健夫・津谷喜一郎編著. 臨床試験と疫学研究のための国際ルール集 (Part 2). ライフサイエンス出版 ;2016. p.148-55.
  - 11) Kamioka H, Tsutani K, Origasa H, Yoshizaki T, Tang W, Shimada M, et al. Quality of systematic reviews of the Foods with Function Claims registered at the Consumer Affairs Agency website in Japan: a prospective systematic review. *Nutrition Research* 2017;40:21-31.
  - 12) Kamioka H, Tsutani K, Origasa H, Yoshizaki T, Shimada M, Kitayuguchi J, et al. Quality of systematic reviews of the Foods with Function Claims in Japan: Comparative before-and after-evaluation of verification reports by the Consumer Affairs Agency. *Nutrients* 2019;11:1583.
  - 13) 上岡洋晴, 津谷喜一郎, 折笠秀樹, 大室弘美, 島田美樹子, 北湯口純ほか. 機能性表示食品制度における届出されたシステマティック・レビューの報告の質 :PRISMA 声明チェックリストに基づく前後比較研究. *薬理と治療* 2019;47:357-73.
  - 14) 消費者庁ホームページ. 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. [http://www.caa.go.jp/foods/pdf/food\\_with\\_function\\_report\\_0001.pdf](http://www.caa.go.jp/foods/pdf/food_with_function_report_0001.pdf)
  - 15) 上岡洋晴, 折笠秀樹編. 機能性表示食品. 適正な研究レビューのための必携マニュアル. ライフサイエンス出版 ;2016.
  - 16) 津谷喜一郎. 食品だから医薬品ではないのだから : “合理的な使用” のなかにを緩めるのか? *薬理と治療* 2018;47:9-12.
  - 17) 上岡洋晴, 山本(前田)万里, 津谷喜一郎, 折笠秀樹, 大室弘美, 島田美樹子ほか. 「PRISMA 声明チェックリスト : 機能性表示食品のための拡張版」における追加項目の開発 : PRISMA FFC extension ver.2. *薬理と治療* 2019;47:1963-73.
  - 18) Page MJ, McKenzie JE, Bossuyt PM, Boutron I, Hoffmann TC, Mulrow CD, et al. The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. *BMJ* 2021;372:n71.
  - 19) 上岡洋晴, 金子善博, 津谷喜一郎, 中山健夫, 折笠秀樹. 「PRISMA 2020 声明 : レビュー報告のための更新版ガイドライン」の解説と日本語訳. *薬理と治療* 2021;49:831-42.
  - 20) Rethlefsen ML, Kirtley S, Waffenschmidt S, Ayala AP, Moher D, Page MJ, et al. PRISMA-S: an extension to the PRISMA statement for reporting literature searches in systematic reviews. *Syst Rev* 2021;10:39.
  - 21) 上岡洋晴, 眞喜志まり, 佐山曉子, 津谷喜一郎, 折笠秀樹. 「PRISMA -S : システマティック・レビュー報告のための PRISMA 声明拡張版」の解説と日本語訳. *薬理と治療* 2021;49:1057-79.
  - 22) Shea BJ, Reeves BC, Wells G, Thuku M, Hamel C, Moran J, et al. AMSTAR 2: a critical appraisal tool for systematic reviews that include randomized or non-randomised studies of healthcare interventions, or both. *BMJ* 2017;358:j4008.
  - 23) 上岡洋晴, 折笠秀樹, 津谷喜一郎. 「AMSTAR 2 : ヘルスケア介入のランダム化または非ランダム化研究あるいは両方を含むシステマティック・レビューのための厳密な評価ツール」の紹介と解説. *薬理と治療* 2018;46:1785-96.
  - 24) Page MJ, McKenzie JE, Bossuyt PM, Boutron I, Hoffmann TC, Mulrow CD, et al. PRISMA 2020 explanation and elaboration: updated guidance and exemplars for reporting systematic reviews. *BMJ* 2021; 372:n160.
  - 25) Chan A-W, Tetzlaff JM, Altman DG, Laupacis A, Gøtzsche PC, Krleža-Jerić K, et al. SPIRIT 2013 Statement: Defining standard protocol items for clinical trials. *Ann Intern Med* 2013; 158: 200-7.
  - 26) 折笠秀樹, 津谷喜一郎, 上岡洋晴. SPIRIT 2013 声明 : 臨床試験のための標準的なプロトコル項目の規定. *薬理と治療* 2017;45:1895-904.
  - 27) 消費者庁ホームページ. 機能性表示食品制度における臨床試験及び安全性の評価内容の実態把握の検証・調査事業報告書. [https://www.caa.go.jp/policies/policy/food\\_labeling/foods\\_with\\_function\\_claims/pdf/foods\\_index\\_23\\_171025\\_0001.pdf](https://www.caa.go.jp/policies/policy/food_labeling/foods_with_function_claims/pdf/foods_index_23_171025_0001.pdf)