

Verification of scientific evidence on effectiveness of the system of
"Foods with Function Claims":
Assessment of the submitted systematic literature reviews
(digest edition)

2016.3
Consumer Affairs Agency

This report is prepared by Mizuho Information & Research Institute, Inc. which is commissioned by the CAA. This report is based on the result of the expert working group organized by Mizuho Information & Research Institute, Inc.

Verification of scientific evidence on effectiveness of the system of "Foods with Function Claims": Assessment of the submitted systematic literature reviews

List of members of the working group

Prof. Hiroharu Kamioka¹

Faculty of Regional Environment Science, Tokyo University of Agriculture

Prof. Hideki Origasa²

Biostatistics and Clinical Epidemiology, University of Toyama School of Medicine

Prof. Hiromi Takano-Ohmuro

Faculty of Pharmacy, Musashino University

Dr. Takahiro Yoshizaki,

Assistant Professor, Faculty of Food and Nutritional Science, Toyo University

Dr. Jun Kitayuguchi

Senior Research Fellow, Physical Education and Medicine Research Center Unnan

Dr. Mikiko Shimada

National Registered Dietitian, Medical Corporation Hidaka Rehabilitation Hospital

Mr. Wentao Tang

Doctor's Course, Department of Drug Policy and Management, Graduate School of Pharmaceutical Sciences, The University of Tokyo

Ms. Satoko Sayama

Librarian, Library, Center for Academic Resources, St. Luke's International University

Ms. Mari Makishi

Librarian, NARASHINO Media Center for Research & Education, Medianet Center, Toho University

¹ Chairperson of the committee

² Vice chairperson of the committee

Chapter 1 Overview of the system and purpose of the assessment

Based on “Committee Report on the new system of foods with function claims issued on July 30, 2014, “the system of foods with function claims” (called “this system” hereinafter) was established in April 2015.

According to the Food Labelling Standards pursuant to the Food Labelling Act, under the responsibility of food business operators, based on scientific evidence of safety and effectiveness, this system allows labelling which indicates that the food is expected to have a specific effect for health, except for reducing risk of diseases, through the process of submission to the Secretary-General of the CAA. (Article 2, paragraph (3), item (1) of the Food Labelling Act). Before this system was in place, making function claims on food labels had only been allowed for Foods for Specified Health Uses and for Foods with Nutrient Function Claims. This system is expected to provide additional choices of products with function claims, and this enables consumers to make informed choices of such products.

According to “Guidelines on Notification of Foods with Function Claims¹” dated March 30, 2015 (called “the guidelines” hereinafter), scientific evidence of effectiveness shall be supported with either clinical trial with a finished product or systematic literature review of a finished product or functional substances. As of October 31, 2015, 122 notifications of foods (including two notifications withdrawn) are published on the website of the Consumer Affairs Agency.

Systematic literature review is a method of study which organizes knowledge obtained from literatures such as research papers, and called “Systematic review” in general (called “SR” hereinafter). Regarding methods of systematic literature review of effectiveness on this system, points of attention are specified in the guidelines. Based on a quantitative or a qualitative systematic literature review performed by a food business operator, only systematic literature reviews in which effectiveness to be claimed can be considered as positive from the viewpoint of totality of evidence may be used as scientific evidence.

In this project, scientific evaluation of systematic literature reviews submitted as evidence of effectiveness of foods was conducted in order to extract issues for appropriate operation of this system for the next year and beyond and to consider measures for improving quality of submitted SRs, the following were conducted:

1. Verification according to the PRISMA Checklist (Chapter 3)
2. Verification of the quality of specific reporting methods in submitted SRs of this system (Chapter 4)
3. Identification of what to improve in submitted SRs for foods with function claims (Chapter 5, Appendix)

¹ “Guidelines on Notification of Foods with Function Claims” Consumer Affairs Agency, 2015.3.30. <http://www.caa.go.jp/foods/pdf/150330_guideline.pdf> (in Japanese) (accessed 2016-03-22).

Chapter 2 Range of matters handled in this report and its role

In this project, submitted SRs were evaluated by considering various evaluation methods in addition to according to the PRISMA checklist. Therefore, it is necessary to clarify the points to be considered in advance to promote correct understanding and interpretation of matters handled in this report and their range. Then, the range of matters handled in this report and its role will be shown in this chapter as below.

The purposes of this verification are to clarify whether a submitted SR has been conducted in accordance with the correct method and to propose how to improve the quality of submitted SRs. Specifically to propose how to describe them by verifying whether there are non-performances, omissions, or uncertainties in the items. Therefore, verification targets refer to the main contexts, figures, and charts described in submitted SRs.

The main aspects of interest of effectiveness claiming for foods are the following three regarding functional substances or finished products:

- Whether they are really safe
- Whether they have clear mechanism of action
- Whether they are really effective

However, those mentioned above cannot be answered unless we go back to “primary studies cited or used as references” described in submitted SRs or other documents and evaluate the quality of each study report.

As they are studies on different levels, this project is limited to evaluation of the quality of reports as mentioned above, not including evaluation of safety, mechanism, and efficacy of functional substances or finished products.

Chapter 3 Verification according to the PRISMA checklist

Section 1 Verification results of submitted SRs based on the “PRISMA Checklist: Extended Version for SR of Foods with Function Claims”

1. Methods

As a result of excluding two withdrawn, 21 notifications based on a clinical trial with a finished product, and 48 submitted SRs of overlapping from 122 submitted SRs made public as of October 31, 2015, 51 submitted SRs were selected for evaluation of the quality (Figure 1).

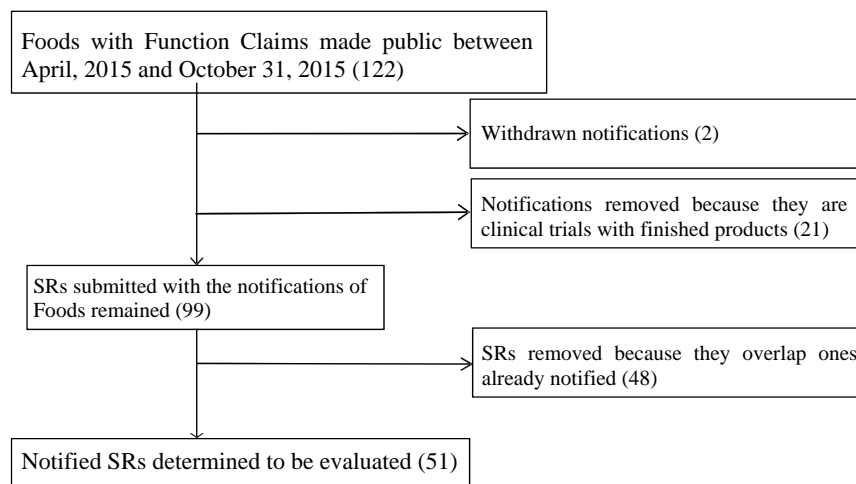


Figure 1. Selection process for submitted SRs to be evaluated.

The list of selected submitted SRs, removed submitted SRs, and reasons for exclusion are shown in the attachments of the full text of this report.

For quality evaluation, we prepared the “PRISMA Checklist: Extended Version for submitted SRs of Foods with Function Claims” by adding 45 low-order items to the “PRISMA Checklist”, which consists of 27 major topics. The PRISMA Checklist often requires multiple entries for one topic and this makes it difficult to determine whether or not each topic is entered appropriately. To solve this problem, we divided one topic into several descriptions and established low-order items based on this system.

2. Results

We defined “non-performance/no description” or “insufficient/unclear description” as a deficiency in the description. Table 1 shows the number of submitted SRs (frequency) and rate (deficiency rate) for all of submitted SRs evaluated in this project. Table 2 shows the number and rate for submitted SRs with meta-analysis, and Table 3 shows those for submitted SRs without meta-analysis. It is necessary to describe items with a high deficiency rate appropriately with special attention given. Please see the main part of the report for points to be improved and detailed description methods obtained from the evaluation results of each item.

Table 1. List of results of quality evaluation based on the “PRISMA Checklist: Extended Version for SR of Foods with Function Claims”

【Submitted SRs (a total of 51 submitted SRs)】

Item number	Evaluation points based on the “PRISMA Checklist: Extended Version for SR of Foods with Function Claims”	Number of submitted SRs with “insufficient/unclear description” (%)	Number of submitted SRs with “non-performance/no description” (%)	Number of deficiencies (%)
#1	Title	0 (0.0%)	0 (0.0%)	0 (0.0%)
#2	Structured summary	33 (64.7%)	0 (0.0%)	33 (64.7%)
#3	Rationale	21 (41.2%)	4 (7.8%)	25 (49.0%)
#4a	Description related to participants	4 (7.8%)	2 (3.9%)	6 (11.8%)
#4b	Description related to interventions	4 (7.8%)	2 (3.9%)	6 (11.8%)
#4c	Description related to comparisons	3 (5.9%)	5 (9.8%)	8 (15.7%)
#4d	Description related to outcomes	4 (7.8%)	2 (3.9%)	6 (11.8%)
#5a	Presence or absence of a review / protocol	9 (17.6%)	17 (33.3%)	26 (51.0%)
#5b	Accessibility to a protocol or not	0 (0.0%)	49 (96.1%)	49 (96.1%)
#5c	Presence or absence of web address and registration number of a protocol	0 (0.0%)	49 (96.1%)	49 (96.1%)
#6a	Description related to study characteristics	16 (31.4%)	1 (2.0%)	17 (33.3%)
#6b	Description related to report characteristics	19 (37.3%)	3 (5.9%)	22 (43.1%)
#7	Information sources	20 (39.2%)	2 (3.9%)	22 (43.1%)
#8	Search	7 (13.7%)	1 (2.0%)	8 (15.7%)
#9a	Description of screening method	21 (41.2%)	0 (0.0%)	21 (41.2%)
#9b	Description of eligibility	12 (23.5%)	0 (0.0%)	12 (23.5%)
#9c	Description of criteria for selection	17 (33.3%)	0 (0.0%)	17 (33.3%)
#10	Data collection process	17 (33.3%)	7 (13.7%)	24 (47.1%)
#11	Data items	12 (23.5%)	1 (2.0%)	13 (25.5%)
#12a	Bias risks	27 (52.9%)	2 (3.9%)	29 (56.9%)
#12b	Indirectness	20 (39.2%)	6 (11.8%)	26 (51.0%)
#12c	Imprecision	12 (25.5%)	4 (8.5%)	16 (34.0%)
#12d	Inconsistency	19 (37.3%)	3 (5.9%)	22 (43.1%)
#13	Summary measures	9 (21.4%)	7 (16.7%)	16 (38.1%)
#14a	Description of method for synthesis of results	2 (22.2%)	2 (22.2%)	4 (44.4%)
#14b	Description of measures of consistency	1 (11.1%)	1 (11.1%)	2 (22.2%)
#15a	Search for clinical trial registration	2 (3.9%)	38 (74.5%)	40 (78.4%)
#15b	Inquiry to authors	1 (2.0%)	46 (90.2%)	47 (92.2%)
#15c	(for post meta-analysis) Funnel plot	0 (0.0%)	1 (11.1%)	1 (11.1%)
#15d	Description of selective report within studies and others	15 (31.9%)	9 (19.1%)	24 (51.1%)
#16	Additional analyses	0 (0.0%)	3 (37.5%)	3 (37.5%)
#17	Study selection	1 (2.0%)	0 (0.0%)	1 (2.0%)
#18	Study characteristics	10 (19.6%)	0 (0.0%)	10 (19.6%)
#19	Risk of bias within studies	13 (25.5%)	1 (2.0%)	14 (27.5%)
#20a	Description of simple summary data on each intervention group	19 (38.8%)	5 (10.2%)	24 (49.0%)
#20b	Forrest plot	1 (11.1%)	1 (11.1%)	2 (22.2%)
#21	Synthesis of results	1 (12.5%)	0 (0.0%)	1 (12.5%)
#22	Risk of bias across studies	17 (33.3%)	0 (0.0%)	17 (33.3%)
#23	Additional analysis	1 (11.1%)	4 (44.4%)	5 (55.6%)
#24	Summary of evidence	31 (64.6%)	8 (16.7%)	39 (81.3%)
#25a	Description of limitations at study and outcome levels	18 (35.3%)	15 (29.4%)	33 (64.7%)
#25b	Description of limitations at review level	20 (39.2%)	10 (19.6%)	30 (58.8%)
#26	Conclusions	29 (56.9%)	10 (19.6%)	39 (76.5%)
#27a	Source for funding of the SR and other support	8 (15.7%)	7 (13.7%)	15 (29.4%)
#27b	Role of funders for the SR	3 (6.1%)	9 (18.4%)	12 (24.5%)

Table 2. List of results of quality evaluation based on the “PRISMA Checklist: Extended Version for SR of Foods with Function Claims”

【Submitted SRs with meta-analysis (a total of 9 submitted SRs)】

Item number	Evaluation points based on the “PRISMA Checklist: Extended Version for SR of Foods with Function Claims”	Number of submitted SRs with “insufficient/unclear description” (%)		Number of submitted SRs with “non-performance/no description” (%)		Number of deficiencies (%)	
#1	Title	0	(0.0%)	0	(0.0%)	0	(0.0%)
#2	Structured summary	6	(66.7%)	0	(0.0%)	6	(66.7%)
#3	Rationale	3	(33.3%)	2	(22.2%)	5	(55.6%)
#4a	Description related to participants	0	(0.0%)	1	(11.1%)	1	(11.1%)
#4b	Description related to interventions	0	(0.0%)	1	(11.1%)	1	(11.1%)
#4c	Description related to comparisons	1	(11.1%)	1	(11.1%)	2	(22.2%)
#4d	Description related to outcomes	0	(0.0%)	0	(0.0%)	0	(0.0%)
#5a	Presence or absence of a review / protocol	1	(11.1%)	4	(44.4%)	5	(55.6%)
#5b	Accessibility to a protocol or not	0	(0.0%)	8	(88.9%)	8	(88.9%)
#5c	Presence or absence of web address and registration number of a protocol	0	(0.0%)	8	(88.9%)	8	(88.9%)
#6a	Description related to study characteristics	2	(22.2%)	1	(11.1%)	3	(33.3%)
#6b	Description related to report characteristics	1	(11.1%)	1	(11.1%)	2	(22.2%)
#7	Information sources	3	(33.3%)	0	(0.0%)	3	(33.3%)
#8	Search	1	(11.1%)	0	(0.0%)	1	(11.1%)
#9a	Description of screening method	4	(44.4%)	0	(0.0%)	4	(44.4%)
#9b	Description of eligibility	1	(11.1%)	0	(0.0%)	1	(11.1%)
#9c	Description of criteria for selection	2	(22.2%)	0	(0.0%)	2	(22.2%)
#10	Data collection process	6	(66.7%)	0	(0.0%)	6	(66.7%)
#11	Data items	0	(0.0%)	1	(11.1%)	1	(11.1%)
#12a	Bias risks	3	(33.3%)	2	(22.2%)	5	(55.6%)
#12b	Indirectness	2	(22.2%)	3	(33.3%)	5	(55.6%)
#12c	Imprecision	1	(11.1%)	1	(11.1%)	2	(22.2%)
#12d	Inconsistency	1	(11.1%)	1	(11.1%)	2	(22.2%)
#13	Summary measures	1	(12.5%)	1	(12.5%)	2	(25.0%)
#14a	Description of method for synthesis of results	2	(22.2%)	2	(22.2%)	4	(44.4%)
#14b	Description of measures of consistency	1	(11.1%)	1	(11.1%)	2	(22.2%)
#15a	Search for clinical trial registration	0	(0.0%)	8	(88.9%)	8	(88.9%)
#15b	Inquiry to authors	0	(0.0%)	7	(77.8%)	7	(77.8%)
#15c	(for post meta-analysis) Funnel plot	0	(0.0%)	1	(11.1%)	1	(11.1%)
#15d	Description of selective report within studies and others	2	(22.2%)	1	(11.1%)	3	(33.3%)
#16	Additional analyses	0	(0.0%)	3	(37.5%)	3	(37.5%)
#17	Study selection	1	(11.1%)	0	(0.0%)	1	(11.1%)
#18	Study characteristics	3	(33.3%)	0	(0.0%)	3	(33.3%)
#19	Risk of bias within studies	2	(22.2%)	0	(0.0%)	2	(22.2%)
#20a	Description of simple summary data on each intervention group	1	(11.1%)	0	(0.0%)	1	(11.1%)
#20b	Forrest plot	1	(11.1%)	1	(11.1%)	2	(22.2%)
#21	Synthesis of results	1	(12.5%)	0	(0.0%)	1	(12.5%)
#22	Risk of bias across studies	3	(33.3%)	0	(0.0%)	3	(33.3%)
#23	Additional analysis	1	(11.1%)	4	(44.4%)	5	(55.6%)
#24	Summary of evidence	6	(66.7%)	1	(11.1%)	7	(77.8%)
#25a	Description of limitations at study and outcome levels	2	(22.2%)	2	(22.2%)	4	(44.4%)
#25b	Description of limitations at review level	3	(33.3%)	2	(22.2%)	5	(55.6%)
#26	Conclusions	3	(33.3%)	4	(44.4%)	7	(77.8%)
#27a	Source for funding of the SR and other support	0	(0.0%)	2	(22.2%)	2	(22.2%)
#27b	Role of funders for the SR	0	(0.0%)	1	(11.1%)	1	(11.1%)

Table 3. List of results of quality evaluation based on the “PRISMA Checklist: Extended Version for SR of Foods with Function Claims”

【Submitted SRs without meta-analysis (a total of 42 submitted SRs)】

Item number	Evaluation points based on the “PRISMA Checklist: Extended Version for SR of Foods with Function Claims”	Number of submitted SRs with “insufficient/ unclear description” (%)		Number of submitted SRs with “non-performance/ no description” (%)		Number of deficiencies (%)	
#1	Title	0	(0.0%)	0	(0.0%)	0	(0.0%)
#2	Structured summary	27	(64.3%)	0	(0.0%)	27	(64.3%)
#3	Rationale	18	(42.9%)	2	(4.8%)	20	(47.6%)
#4a	Description related to participants	4	(9.5%)	1	(2.4%)	5	(11.9%)
#4b	Description related to interventions	4	(9.5%)	1	(2.4%)	5	(11.9%)
#4c	Description related to comparisons	2	(4.8%)	4	(9.5%)	6	(14.3%)
#4d	Description related to outcomes	4	(9.5%)	2	(4.8%)	6	(14.3%)
#5a	Presence or absence of a review / protocol	8	(19.0%)	13	(31.0%)	21	(50.0%)
#5b	Accessibility to a protocol or not	0	(0.0%)	41	(97.6%)	41	(97.6%)
#5c	Presence or absence of web address and registration number of a protocol	0	(0.0%)	41	(97.6%)	41	(97.6%)
#6a	Description related to study characteristics	14	(33.3%)	0	(0.0%)	14	(33.3%)
#6b	Description related to report characteristics	18	(42.9%)	2	(4.8%)	20	(47.6%)
#7	Information sources	17	(40.5%)	2	(4.8%)	19	(45.2%)
#8	Search	6	(14.3%)	1	(2.4%)	7	(16.7%)
#9a	Description of screening method	17	(40.5%)	0	(0.0%)	17	(40.5%)
#9b	Description of eligibility	11	(26.2%)	0	(0.0%)	11	(26.2%)
#9c	Description of criteria for selection	15	(35.7%)	0	(0.0%)	15	(35.7%)
#10	Data collection process	11	(26.2%)	7	(16.7%)	18	(42.9%)
#11	Data items	12	(28.6%)	0	(0.0%)	12	(28.6%)
#12a	Bias risks	24	(57.1%)	0	(0.0%)	24	(57.1%)
#12b	Indirectness	18	(42.9%)	3	(7.1%)	21	(50.0%)
#12c	Imprecision	11	(28.9%)	3	(7.9%)	14	(36.8%)
#12d	Inconsistency	18	(42.9%)	2	(4.8%)	20	(47.6%)
#13	Summary measures	8	(23.5%)	6	(17.6%)	14	(41.2%)
#14a	Description of method for synthesis of results	-	-	-	-	-	-
#14b	Description of measures of consistency	-	-	-	-	-	-
#15a	Search for clinical trial registration	2	(4.8%)	30	(71.4%)	32	(76.2%)
#15b	Inquiry to authors	1	(2.4%)	39	(92.9%)	40	(95.2%)
#15c	(for post meta-analysis) Funnel plot	-	-	-	-	-	-
#15d	Description of selective report within studies and others	13	(34.2%)	8	(21.1%)	21	(55.3%)
#16	Additional analyses	-	-	-	-	-	-
#17	Study selection	0	(0.0%)	0	(0.0%)	0	(0.0%)
#18	Study characteristics	7	(16.7%)	0	(0.0%)	7	(16.7%)
#19	Risk of bias within studies	11	(26.2%)	1	(2.4%)	12	(28.6%)
#20a	Description of simple summary data on each intervention group	18	(45.0%)	5	(12.5%)	23	(57.5%)
#20b	Forrest plot	-	-	-	-	-	-
#21	Synthesis of results	-	-	-	-	-	-
#22	Bias risks across studies	14	(33.3%)	0	(0.0%)	14	(33.3%)
#23	Additional analysis	-	-	-	-	-	-
#24	Summary of evidence	25	(64.1%)	7	(17.9%)	32	(82.1%)
#25a	Description of limitations at study and outcome levels	16	(38.1%)	13	(31.0%)	29	(69.0%)
#25b	Description of limitations at review level	17	(40.5%)	8	(19.0%)	25	(59.5%)
#26	Conclusions	26	(61.9%)	6	(14.3%)	32	(76.2%)
#27a	Source for funding of the SR and other support	8	(19.0%)	5	(11.9%)	13	(31.0%)
#27b	Role of funders for the SR	3	(7.5%)	8	(20.0%)	11	(27.5%)

Chapter 4 Verification of the quality of specific reporting methods in submitted SRs of the system of “Foods with function claims”

For 51 submitted SRs extracted by the process in Figure 1, we verified the following: “appropriateness related to search”, “appropriateness of evaluation of bias risks for individual study”, “appropriateness of evaluation of evidence as a whole”, “appropriateness of the procedure and description method of the meta-analysis”, and “appropriateness of other specific contents”.

Evaluation results of each item and issues to be considered are described below.

Section 1 Verification of appropriateness related to literature search

For 51 submitted SRs extracted by the process in Figure 1, appropriateness related to literature search was verified. If correction or revision was added to SRs during verification, the corrected or revised submitted SRs were verified.

For verification of appropriateness, “Description of Basic Items Related to Search: For submitted SRs of Foods with Function Claims” was prepared and the number and rate of submitted SRs corresponding to each item were calculated. Moreover, “the search formula (search strategy)” was evaluated using “evaluation items for the search formula”.

Evaluation results and issues are as follows.

1. Description of basic items related to search

(1) Appropriateness of the searcher

In more than half of the submitted SRs, abilities and experiences of the searchers were not described. The guidelines do not require search by a literature search specialist including an information search technician or a librarian.

However, a comprehensive search using various databases may play an important role in improving the quality of submitted SRs and require expertness, so the participation of the literature search specialist is recommended.

(2) Entry of the search date

Although it is essential to enter the correct search date, some of the submitted SRs were significantly old. To conduct SRs, it is also important to collect the latest studies.

(3) Sources to search

Almost all of the submitted SRs used PubMed, and they also used various kinds of databases such as JDream II and ICHUSHI Web (database of medical literature written in Japanese). In the guidelines, types of sources including bibliographic databases for searching are not specified, but it is provided that “sources shall be selected appropriately among those regarded as objectively relevant for a literature search in the field”. Additionally, it is pointed out that

articles not contained in PubMed exist in some fields. In fact, to collect comprehensive evidence including SRs, it is important to search as many databases and clinical trial registration systems as possible.

(4) Entry of search formula

The PRISMA checklist (2009), with which compliance is required in the guidelines, provides that “Present full electronic search strategy for at least one database, including any limits used, such that it could be repeated.” However, in some submitted SRs, the formulas were entered unclearly or the number of search results was not reproduced by the entered formula. It is essential to enter not only the search formula but also the number of search results correctly.

2. Evaluation of search formula (search strategy)

The results of evaluation for appropriateness of search formula showed differences among submitted SRs. This may be mainly attributable to the searchers’ literature search ability. In some submitted SRs, searches were conducted by combining keywords and thesauruses appropriately and setting the optimal search formula according to the database characteristics, while in other submitted SRs, the search formulas were not entered correctly due to lack of search keywords or unnecessary limitations by year or outcomes.

The guidelines define the following: “To search comprehensively, a search formula made by combining free items and controlled terms (including MeSH for PubMed) appropriately will be set per bibliographic database” and “From the viewpoint of removing a language bias, articles written in English and Japanese at least shall be searched”. To search for SRs, it is important to set the optimal search formula made by combining keywords and thesauruses (such as MeSH) appropriately for each CQ (Clinical Question) according to the database characteristics. Moreover, it is desirable not to include O of PI(E)CO² in the search formula, because a comprehensive search for SRs and clinical studies using P, I(E), and (C) of PI(E)CO may lead to various outcomes including harm and disadvantage. Basically, it is desirable not to narrow the search results with a filter including language, even if the number of articles is large. However, if stated in the protocol in advance, a narrowing search is expected to be conducted depending on the workload.

Section 2 Appropriateness of evaluation of bias risks for individual study

Appropriateness of bias risks for individual study used for submitted SRs was evaluated. The results showed insufficient description of the method of evaluation of bias risks. Lack of description of the evaluation method makes it impossible for a third party to not only reproduce submitted SRs but also determine the obtained result objectively. In particular, description of

² The acronym for Participants, Interventions (or Exposures), Comparators and Outcomes.

whether the bias risks were evaluated by two independent persons reflects how carefully and precisely the target article was examined. It is supposed to describe the degree of agreement between the two reviewers or coefficient κ ; however, these descriptions were found in very few submitted SRs to be evaluated. Additionally, descriptions related to bias risks were found in approximately 90% of submitted SRs; however, the majority of them were “as described in the attachment” or similar expressions.

The evaluation of bias risks shall be described in the “Methods and Results” and they should be carefully considered in the “Discussion” to clarify how the evaluation results are reflected in the conclusion of submitted SRs, but it was not described specifically in most submitted SRs. With regard to evaluation of bias risks for individual study, it is desirable to describe how the evaluation results were reflected to the criteria for selection of submitted SR target articles, including by what level of bias risks was selected, in addition to evaluation methods.

Section 3 Verification of appropriateness of evaluation of evidence in all (Indirectness, Imprecision, and Inconsistency)

Individual evaluation methods for Indirectness, Imprecision, and Inconsistency were well described in only about 20% of submitted SRs. Evaluation results were described in only about 20-40% of them and it is difficult to say that many submitted SRs allow a third party to grasp and understand the content of evaluation of evidence. Regarding all the items of bias risk, Indirectness, Imprecision, and Inconsistency, both methods and results were well described in only four submitted SRs (7.8%). This result shows that many submitted SRs lack sufficient information from which a conclusion can be determined objectively. In one case where there was only one article targeted for submitted SR, the evaluation results of Imprecision and Inconsistency were described in the main part of the article or Form (V)-13a (“Form” omitted hereinafter) and there remained questions about the evaluation method. Therefore, it is expected to ensure more transparency for evaluation methods of Indirectness, Imprecision, and Inconsistency. Additionally, less than 30 % of the submitted SRs in which how the evaluation of Indirectness, Imprecision, and Inconsistency was reflected in the evaluation of evidence in all were clearly described. Evaluation contents including this point should be described in the main part.

Section 4 Verification of appropriateness of procedures and description of meta-analysis

The sufficiency rate of procedures and descriptions for meta-analysis was high for almost all items of evaluation. This may be attributable to the use of statistical software for meta-analysis by an analyzer. However, some cases were identified in which the rate was manually calculated not using any statistical software. The meta-analysis methods of 6 out of 9 submitted SRs were a random-effects model (DerSimonian-Laird method), which may be attributable to performance by the same analyzer.

As an issue to be considered, it will be required to describe how studies with a lack of information about mean/standard deviation are handled. In particular, some articles used for meta-analysis may have a description of mean but no information about standard deviation. For these articles, it is desirable to describe whether they existed or not and how they were handled (e.g., they were calculated by 95% CI (Confidence Interval), they were excluded, or they were made up for with mean differences). A meta-analysis should basically integrate only RCTs; if not, it will be required to perform further sensitivity analysis (even if only RCTs are integrated). If no consistency is identified, or if publication bias or instability of the result in a sensitivity analysis is suspected, although none of them was found this time, it is expected to perform a meta-analysis with more careful attention.

When a meta-analysis, or a quantitative SR, is performed, it is desirable to describe the eight items shown in Table 3.

Table 4. Eight items that should be included in meta-analysis reporting

- 1) Index to be integrated (mean difference, odds ratio, etc.)
- 2) Integration method (random-effects model, etc.) and weight applied to the study
- 3) Description of Forest plot
- 4) Consideration of heterogeneity (Cochran's Q statistics, I^2 statistics, etc.)
- 5) Consideration of publication bias (Funnel plot, etc.)
- 6) Sensitivity analysis (subgroup analysis, etc.)
- 7) Statistical software used
- 8) Outsourcing operators

Section 5 Verification of appropriateness of other specific contents in submitted SRs of this system

1. Reviewer's characteristics

Some of the submitted SRs, participations of a librarian, a degree recipient, or an EBM expert were not described. The guidelines do not require them to participate in the review. However, those who have expert knowledge will play an important role in comprehensive searching for articles using various databases and reviewing searched articles in order to improve the quality of submitted SRs, and it is recommended.

On the other hand, their responsibility is huge since they are experts, and this matter is related to conflict of interest. Therefore, it is preferable to clarify their names and the organizations to which they belong for their participation.

2. Conducting a hand search

12 submitted SRs (23.5%) did not describe whether a hand search was carried out or not. A hand search is an easy method to perform using academic journals or abstracts for professional meetings and it is essential for searching. When using this method, it is necessary to definitely describe the volume, issue number, date of publication of academic journal, etc. as well as the method.

3. Independence of those in charge of screening

In more than half of submitted SRs, more than two people were participated in screening of articles. Basically, an SR is a study conducted by multiple persons. This is based on the premise that "every human can make a careless mistake" and it indicates the importance of double-checking. At the phase of screening articles, if those that meet eligibility criteria are missed, the phase goes from screening to analysis without missed articles being noticed, and, ultimately, it is highly possible to conduct an SR incorrectly. Therefore, it should be essential to conduct screening independently by more than two people.

4. Selected study design

In 14 submitted SRs (27.5%), the study design was not described or unclear. It is important to describe the methods using the PI(E)COS³ format and especially essential to describe the study design to be selected in the main part of the article. Otherwise, there is a risk that a study design that can bring a biased result may be selected. As related to preliminary registration of SRs, it is necessary to conduct a literature search after a study design is definitely determined.

³ The word added "Study design (S)" to PI(E)CO.

5. Selection of data on participants with disease

A few submitted SRs selected articles including participants beyond the range of mild disease. The purpose of this system is to help improve health and not to treat or prevent disease. Data on participants with disease cannot be used for submitted SRs. There is a scientific premise that if participants' initial values in the primary or secondary outcomes are significantly high or low, it indicates that study data on participants with disease can be included. Such study data should not be used for submitted SR.

The guidelines allow for the handling of participants with mild disease within a certain range; however, it is necessary to include participants without disease in the target population and conduct SR using data on only participants without disease, and the result will be reported in both a study review report and an abstract for a study review for consumers. Therefore, each data should be described separately in the "Results" section of the main part, (V)-11 and (V)-13 at least, and discussed comprehensively. Additionally, if sources used as references include data on participants with disease, it is essential to separate the sources from those used for submitted SRs.

In any case, the use of data whose belonging to participants with or without disease is difficult to determine may complicate the situation, and such data should not be included in the study plan for submitted SRs to the extent possible.

Chapter 5 Improvement of submitted SRs for foods with function claims

In this chapter, considering the verification results mentioned in the previous chapters, we show detailed points for SRs and examples as well as what should not be done (items related to inappropriateness and research ethics and items requiring attention).

Section 1 Policy of appropriate description and precautions

From the viewpoint of the food business operators, the issue of a great concern should be “how to describe understood verification results and theories”. Then, we show examples of appropriate description in the attachment based on the “PRISMA Checklist: Extended Version for submitted SRs of Foods with Function Claims”. For covering almost all submitted SRs regardless whether they are qualitative or quantitative, examples are separately shown by submitted SRs with and without meta-analysis. The attachment shows examples for improving the quality of SRs, and it will be useful for preparing SRs and the other documents.

Some of the submitted SRs in this verification, item numbers and names of the PRISMA Checklist are described in the main part of the article (e.g., #13, summary of measures). In the case of SRs as general academic papers, it is common practice to enter the page number on which an item is described in the PRISMA Checklist attached at submission. However, description of item numbers and names in the main part of the article enables us to find points to be reported clearly, and this description should be recommended for submitted SRs for foods with function claims. Therefore, we would like to recommend the entry of item numbers and names used for the PRISMA Checklist in the main part of the article. This will definitely enhance the quality of reporting.

In addition, it is important to make an appropriate study plan for appropriate notification regardless of whether the study is a clinical trial or an SR, and the quality of the study can be determined by the study plan.

Section 2 Precautions considered inappropriate and against research ethics

There are items that cannot be done or must be done regardless of intent, in order to prevent inappropriateness and to comply with research ethics. See the full text of this report for several issues that may arise in submitted SRs and methods for dealing with them.

Section 3 Other precautions

We summarized other precautions especially for describing contents regarding function claims. The most important thing for function claims is to reflect precisely the findings obtained based on PI (E) CO. The basis of the idea is as follows: P (participant) is defined as “what type of target population can be defined based on gender, age, or physical characteristics obtained from reviewed primary studies (it is essential to determine whether the population is limited or not,

e.g., group of only women or only young or aged persons); I (intervention: investigational substance) and E (exposure) describe the “dose that can bring efficacy, time required for occurrence of effects, and dose-response relationship”, while C (comparison control) describes “whether a food or component is appropriate for comparison”. And O (outcome) defines “whether evaluation items are consistent with the contents of claims, whether it is possible to use the description of the contents of claims for the evaluation items, whether outcomes are interpreted in an extensive fashion, whether significant outcomes are obtained by coinciding with multiple hypotheses etc., in which part of the body the outcome is obtained (whether the outcome is misunderstood as an evaluation result for the whole body or if it is for one part of the body)”.

These are basic items and we shall not forget the premise that submitted SRs should be obtained from a certain level or higher of primary study results that can be strong evidence, not by adding SRs later but conducting them under a study plan carefully prepared in advance, collecting all of primary study data including unfavorable results or removing primary studies with high bias. For the above reason, claims are required not only to give consumers universal and correct information but also to prevent consumers misunderstanding information.

Chapter 6 Conclusion

In this report, issues of submitted SRs were discovered by analyzing the current situation and made a proposal to conduct SRs appropriately.

As shown by Tsutani (2003)⁴, “Evidence has the three aspects of creating, informing, and using”. Applying the idea to this system, “creating” corresponds to “primary studies”, while “informing” corresponds to submitted SRs and “using” corresponds to administrators in charge of food, food companies, and consumers. The significance of submitted SRs is “to notify results of primary studies properly”. We consider that a person responsible for notification can conduct an SR that includes minimum information, less bias, and a certain level of quality by using this report as a reference to appropriately make a study plan, conduct a study, and describe results. It is not necessary to achieve high-level SRs such as those conducted in academic researches. To complete basic standard-level SRs, considering the difficulty in handling foods, we attached “appropriate description in SRs based on the ‘PRISMA Checklist: Extended Version for submitted SRs of Foods with Function Claims’” to the report and showed detailed proposals and examples. The attachment will help food business operators who are concerning or questioning about the outline and specifics of SRs.

In conclusion, it has been one year since the system started, that is, it can be also said to be as young as one year old in human age and this system is still developing. Through various constructive discussions between those concerned, continued educational support by academic

⁴ Tsutani.K. “Examination of evidences on EBM” *Therapeutic Research*, 2003; 24:1415-1422. (in Japanese)

institutions, and careful activities to provide consumers with information by those who are responsible for notification (including outsourcing operators of SR) with high ethical values (cooperate ethics and research ethics).