

This is the authors' final, peer reviewed manuscript published in
Journal of Clinical Epidemiology 2010 Jun;63(6):620-26 with the same title.
<http://www.jclinepi.com/article/S0895-4356%2810%2900028-4/abstract>

Summary of Findings tables in Cochrane reviews improved understanding and rapid retrieval of key information.

Sarah E Rosenbaum¹, Claire Glenton², Andrew D Oxman¹

¹Norwegian Knowledge Centre for the Health Services

²Department of Global Health, Sintef, Oslo Norway

*Corresponding author. Email: sarah@rosenbaum.no

Accepted for publication in

Journal of Clinical Epidemiology

22 December 2009

Published: May 2010

Abstract

Objective: To measure the effects of a Summary of Findings (SoF) table on user satisfaction, understanding and time spent finding key results in a Cochrane review.

Study design and setting: We randomized participants in an evidence-based practice workshop (RCT I) and a Cochrane Collaboration entities meeting (RCT II), to receive a Cochrane review with or without a SoF table. In RCT I we measured user satisfaction. In RCT II we measured correct comprehension and time spent finding key results.

Results: RCT I: Participants with the SoF table (n=47) were more likely to “Agree” or “Strongly agree” that it was easy to find results for important outcomes than (n=25) participants without the SoF table: 68% versus 40% (p=0.021). RCT II: Participants with the SoF table (n=18) were more likely to correctly answer two questions regarding results than (n=15) participants without the SoF table: 93% versus 44%(p=0.003) and 87% versus 11%(p<0.001). Participants with the SoF table spent an average of 90 seconds to find key information compared to four minutes for participants without the SoF table (p=0.002).

Conclusion: In two small trials we found that inclusion of a SoF table in a review improve understanding and rapid retrieval of key findings compared to reviews with no SoF table.

Keywords

Knowledge translation; health numeracy; risk communication; systematic reviews; information design; randomized control trial

What's new?

Key finding: Inclusion of a SoF table in a Cochrane review helped readers understand the results more correctly and find key information about the main outcomes and quality of the evidence faster compared to a review with no table.

What this adds to what is known? Evidence summaries and abstracts of systematic reviews may be improved by adding SoF tables.

What are the implications, what should change now? SoF tables will be included in Cochrane reviews. Authors and publishers of other systematic reviews and evidence summaries should consider including SoF tables to facilitate more effective and efficient uptake of key information.

BACKGROUND

Summaries of evidence for health professionals exist in different formats and for different purposes[1]. Structured abstracts, sometimes the only part of a study or review that readers view or use[2, 3], were originally developed to assist readers in retrieving, selecting and critically appraising relevant literature[4-6]. More recently, other forms of summaries have surfaced, such as the ELPS (Electronic long, paper short), Short Cut and Pico formats developed by BMJ[7, 8], motivated by the need for a better utilization of the respective advantages of paper versus the web. Secondary journals, such as ACP (American College of Physicians) Journal Club and other evidence-based journals [9Last accessed: October 29, 2008., 10-12], produce brief summaries of individual studies and reviews, selected for their clinical relevance and newsworthiness. These quality-assessed resources aim to limit the number of journals one needs to access in order to keep abreast with new research and the amount of effort needed to spend on critical appraisal. Online services, such as Clinical Evidence[13] and UpToDate[14], package evidence summaries together with general facts about the topic, recommendations or links to guidelines and other references to create a comprehensive one-stop reference.

The format, content, and representation of the data in these summary types vary. While some include tables or figures, most tend to be dominated by text. Although there is a growing amount of literature in the area of risk communication for consumers[15-18], we found few published studies specifically evaluating different forms of abstracts, synopses or summaries with regards to their effect on clinicians' understanding of the main messages or their decision making[19, 20].

We sought to develop and evaluate a summary format that could be understood by users of systematic reviews, both health professionals and other relevant groups, and that was feasible to implement in Cochrane reviews. The project is one of several

initiatives within the Cochrane Collaboration[21] in recent years to develop and evaluate summaries of Cochrane reviews for different target groups.

Our starting point was the GRADE Summary of Findings (SoF) table. GRADE stands for "Grading of Recommendations Assessment, Development and Evaluation" and is a system for evaluating quality of evidence, and encourages authors to report on the most important outcomes, including adverse effects[22, 23]. An output from GRADE is a table presenting the results (or lack of them) and quality of evidence scores on a single page in a standardized format. Based on extensive feedback from stakeholders and testing with users, we developed a SoF table for inclusion in Cochrane Reviews. The development of the SoF table's content, formatting and data representation is described in a separate article[24].

Here we report our evaluation of the SoF table. We measured the effect of including this SoF table in a Cochrane review on user satisfaction, understanding and time spent finding key results.

Methods

We carried out two randomized controlled trials (RCT's) at different stages of the SoF table development. The objective of the first RCT was to assess users' *satisfaction* with the SoF table; the second RCT aimed to assess the effect of the SoF table on users' *understanding* of the reviews and *time spent* to find answers.

Participants

The first trial took place during a plenary session at workshop for newcomers to evidence-based practice. Participants were asked if they would help evaluate ways of making reviews more accessible, and were told they did not need to participate or return the questionnaires. Seventy-two workshop participants and tutors completed the questionnaires out of a total of approximately 90 people present. These were largely

health professionals, many of whom were ‘beginners’ in evidence-based health care and did not have English as their first language.

The second RCT took place at a meeting for members of Continental European Cochrane entities. The 33 participants were staff from Cochrane entities, including review group coordinators, trial search coordinators and Centre staff. Everybody present participated. Most did not have English as their first language.

Participants in both RCTs had at least a basic understanding of what a systematic review is.

Comparisons

Participants were randomized to three groups:

- the review *without* SoF table,
- the review *with* SoF table (placed after the abstract) with limited formatting (as would be possible in current Cochrane review SoFtware), or
- the review *with* SoF table (placed after the abstract) with full formatting (as might not be possible in current Cochrane review SoFtware).

Randomization

We used block-randomization with 25 blocks of three that was generated on <http://www.randomization.com>. The questionnaires were numbered sequentially the day before and were passed out to all of the participants at each meeting.

Outcome measurement

We measured outcomes through a multiple-choice questionnaire that included questions about the participant and their degree of satisfaction with the accessibility of the main findings of the review. Participants first answered the questionnaire based on the version of the review they had received. Then all participants were shown both formatting versions of the SoF tables and were instructed to answer a final set of

questions measuring their preferences and attitudes about the inclusion Summary of Findings tables in reviews.

In the second trial, the questionnaire was modified to include questions that measured actual understanding. We also asked participants to note how long it took them to find information.

Structured discussions were carried out at the end of both trials.

Sample size

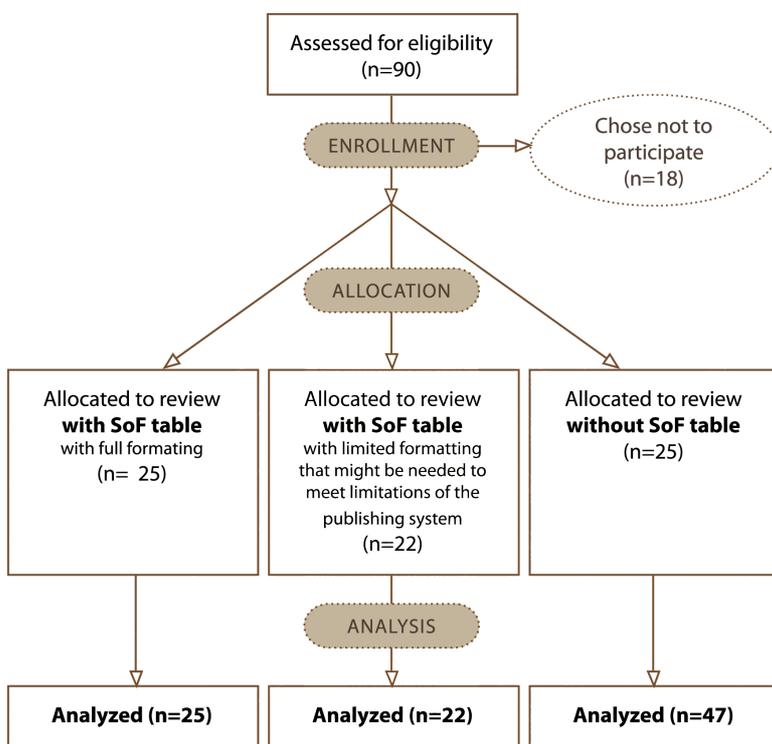
We used convenience samples for the two trials reported here. We initially planned to estimate the sample size for a larger trial based on an alpha of 0.05 for the overall perceived accessibility of the main findings of the review and 80% power to detect a difference of one in the mean rating of accessibility, using data from these studies to estimate the expected mean in the control groups and standard deviations, without adjustment for clustering (within settings) or for multiple comparisons. Secondary analyses of the other questions in the questionnaire were intended primarily to help explain and interpret the results of the primary analysis. Following the first trial we elected to focus the primary analysis on a more objective outcome measure: the proportions of participants answering correctly questions about the risks of symptomless deep vein thrombosis (DVT) with and without compression stockings for people at low risk. Assuming 50% of participants would answer correctly without the SoF table, an alpha of 0.05 and 80% power we would need 58 participants per group to detect a 50% relative improvement. However, due to time and resource constraints and the magnitude of the effect observed in the second trial we did not recruit further participants.

RESULTS

RCT I – Assessing user satisfaction with the SoF table

Of approximately 90 people present, 72 completed the questionnaires (25 without Summary of Findings table, 25 with formatted Summary of Findings table and 22 with Summary of Findings table with limited formatting that might be needed to meet limitations of the publishing system). See Figure 1 for flow diagram.

Figure 1. Flow diagram RCT 1



We have merged the data for the limited and full formatting versions, as the results showed no significant differences between these two versions. For information about the participants, see Table 1.

Table 1. Information about participants, RCT 1

Total n=72 <i>of which 51 health professionals and 28 researchers (11 marked both, 4 marked neither)</i>	Never (before now)	Less than once per year	1 to 4 times per year	5 to 10 times per year	More than 10 times per year
How often do you read reports of randomized trials?	6	4	20	3	39
How often do you read systematic reviews?	7	5	23	12	20
How often do you use the Cochrane Library?	23	5	10	7	22

	Very difficult	Difficult	Easy	Very easy
How easy is it for you to understand reports of randomized trials?	2	29	34	5
How easy is it for you to understand reports of randomized trials?	1	34	32	3

Participants reading reviews that included the Summary of Findings table were more likely to respond that:

- The results and quality of evidence were easy or very easy to find and to understand
- The main findings were easy to understand
- The main findings were accessible.

These differences ranged from 12 to 28% and were not statistically significant except for two measurements: the ease to find results for important outcomes and the perceived accessibility of the quality of the evidence. See Table 3.

Table 3. Perceived understanding and ease of use, RTC 1

What was measured (agreement with statements, on an anchored scale of 0 to 7, where 0 = strongly disagree and 7 = strongly agree.) 6= I agree 7= I strongly agree	% replying “Agree” or “Strongly agree” % of participants answering 6 or 7, combined (CI = 95%)		
	Without SoF table (n=25)	With SoF table (both format versions, n=47)	p-value Pearson Chi-square
Review authors have indicated what they believe are the most important outcomes	80% (64-96%)	81% (70-92%)	p=0.93
It was easy to find this information (about which outcomes are important)	60% (41-79%)	72% (60-85%)	p=0.29
It was easy to find results for important outcomes	40% (21-59%)	68% (55-81%)	p=0.021
It was easy to find quality of the evidence for important outcomes	24% (7-41%)	43% (28-57%)	p=0.12
It was easy to understand the main findings	56% (37-75%)	68% (55-81%)	p=0.31
Main findings are presented in a such a way that they would be helpful to me in making a decision	52% (32-72%)	60% (46-74%)	p=0.54
Overall perceived accessibility Scale 0-7, where 0 is “very inaccessible” and 7 is “Very accessible”	17% (2-32%)	41% (27-56%)	p=0.037

Eighty-one percent agreed or strongly agreed that Cochrane reviews should include SoF tables, 65% with the proposed format, and 75% found the explanation sheet helpful.

Structured discussion

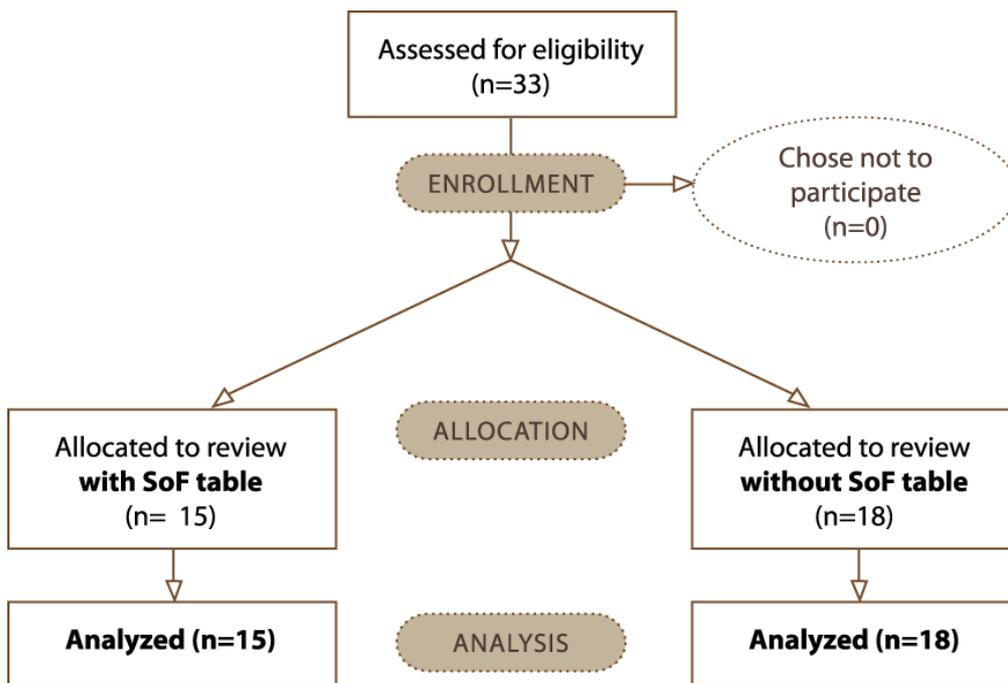
Although most people felt the results were accessible with or without the Summary of Findings tables, many people gave the wrong answer when they were asked to calculate the intervention group risk in a structured discussion following the trial. This discussion led us to the same finding we uncovered earlier through user tests of the SoF table (insert/repeat article reference # 24): presenting results of the intervention group risk framed as an absolute difference (9 fewer per 1,000) caused comprehension problems or uncertainty.

RCT II - Assessing the effect of the SoF table on users' correct understanding of reviews and time spent finding answers

After the first RCT, we made a major change based on the finding that many participants had misunderstood the numbers expressing absolute effect: we replaced this column with a column for intervention group risk.

In the second RCT, using a revised version of the SoF table, we tested correct understanding of the tables and time spent to find answers. Thirty-three completed questionnaires were returned (18 without and 15 with an SoF table). See flow diagram Figure 2.

Figure 2. Flow diagram, RTC II



Participants were staff from Cochrane Entities who held editorial, administrative, technical or marketing positions within the Cochrane Collaboration. Six reported a health professional background and 13 reported a background in research. For information about the participants, see Table 2.

Table 2. Information about participants, RCT II

Total n=33 (of which 6 were health professionals and 13 were researchers)	Never (before now)	Less than once per year	1 to 4 times per year	5 to 10 times per year	More than 10 times per year
How often do you read reports of randomized trials?	3	4	5	4	16
How often do you read systematic reviews?	3	1	8	5	15
How often do you use the Cochrane Library?	2	0	1	1	28

	Very difficult	Difficult	Easy	Very easy
How easy is it for you to understand reports of randomized trials?	1	17	10	1
How easy is it for you to understand reports of randomized trials?	1	18	11	1

The results are shown in Tables 4 and 5. There were large differences in the proportion that answered correctly questions about the risk in the control group (44 versus 93%, $p=0.003$) and the risk in the intervention group (11 versus 87%, $p<0.001$). There were also large differences in actual time spent as well as the proportion that agreed or strongly agreed that it was easy to find information about the quality of evidence for the main outcomes (24 versus 73%, $p=0.005$, Pearson's Chi-square).

Most participants were positive about including the SoF tables and the format: 88% agreed or strongly agreed that Cochrane reviews should have Summary of Findings tables, 84% with the format proposed, and 77% that the explanations were helpful. Most (67%) preferred placing Summary of Findings tables after the abstract.

Table 4. Correct understanding, RCT II

Questions asked	% Correct answer (CI=95%)		p-value Pearson Chi-square
	Without SoF table (n=18)	With SoF table (n=15)	
What is the risk of symptom-less DVT following a long flight for people at low risk who do not wear compression stockings?	44% (21-67%)	93% (81-100%)	p=0.003
What would be the risk if they wore stockings?	11% (0-26%)	87% (69-100%)	P<0.001
How much confidence do the review authors have in the estimated effect of stockings on the risk of symptom-less DVT?	67% (45-88%)	87% (69-100%)	p=0.18
What are the most important outcomes?	33% (9-57%) answered all 6 correctly	53% (28-79%) answered all 6 correctly	p=0.27

Table 5. Correct understanding, RCT II

Questions asked	Mean time spent finding answer Without SoF table (n=18)	Mean time spent finding answer With SoF table (n=15)	p-value Mann-Whitney U-test
	What is the risk of symptom-less deep vein thrombosis (DVT) following a long flight for people at low risk who do not wear compression stockings?	4 minutes	
What would be the risk if they wore stockings?	2.8 minutes	1.3 minutes	p=0.118
How much confidence do the review authors have in the estimated effect of stockings on the risk of symptom-less DVT?	1.5 minutes	2.1 minutes	p=0.47
What are the most important outcomes?	1.9 minutes	2.0 minutes	p=0.88
Would you have needed more time to find this information?	64% would <u>not</u> need more time	79% would <u>not</u> need more time	

DISCUSSION

Results from these evaluations showed that inclusion of the final version of a SoF table in a Cochrane Review helped readers understand the review correctly and made it easier to find information about the quality of evidence and the main outcomes, compared to a review with no SoF table.

This study has some limitations:

- In the first RCT we tested people's satisfaction with the SoF tables rather than actual effect on comprehension. However, we redesigned the protocol for the second RCT in order to measure correct comprehension. For this reason the two RCT's are not identical and cannot be directly compared.
- The RCT's were small.
- Although the second RCT included health professionals and other users of systematic reviews, this group was not necessarily representative of clinicians.
- The formatting of the SoF table we tested may not be possible to achieve in the Cochrane Library for technical reasons. It is not clear how much of a difference that will make.

User satisfaction

Participants for the most part perceived the review with the SoF table as more accessible. But user testing conducted just prior to the first trial and structured discussions conducted after this trial revealed that many had actually misunderstood content in the SoF table[24]. User satisfaction, though important because it relates to motivational issues, can be a misleading indicator of the success of a piece of information if used alone. A systematic review looking at the effect of different graph designs, for instance, found several studies where versions that users preferred performed poorly when the same participants' decision making behaviour was measured[25].

Time spent

Lack of time is an often-cited barrier to uptake of evidence by clinicians. A literature review from 2007 found studies reporting varying amounts of time needed to search for information, ranging from 53 minutes to medians of 14.4 and 15 minutes[26]. One study found that primary care physicians spent on average less than two minutes pursuing an answer to a clinical question[27]. If this is the case, then the findings from our study may be important: time spent to find key results dropped from an average of 4 minutes and 2.8 minutes to 1.5 and 1.3 minutes respectively. This could affect clinicians' ability (or willingness) to access the evidence within timeframes that are realistic in clinical contexts.

Comprehension issues – poor skills or poor information design?

Health professionals may be experts in their medical field but many are non-experts when it comes to biostatistics. Studies have consistently shown for more than 25 years that health care providers display poor health numeracy skills such as basic computation, estimation and statistical literacy[3, 28-33]. High level of education is not a vaccination against low numeracy skills[34]. However a health numeracy model by Ancker and Kaufman, based on a theory of distributed cognition, emphasizes the interdependence of the skills of the care provider with the characteristics of the information[35]. An appropriately designed piece of information can compensate for lack of skills in many ways, e.g. by easing cognitive load, making computation simpler, filling in background knowledge the user doesn't have, or rendering this background knowledge unnecessary.

This was illustrated clearly through both the developmental and evaluations phases of this project. First we observed how small design details made a difference to participants' experience of the table in user testing, then demonstrated in an RCT that the inclusion of a SoF table in a systematic review improved the participants' correct understanding of the key results. The participants in the second RCT were all involved

in the work of the Cochrane Collaboration, and could therefore be expected pick up information from the original Cochrane Review quickly. This makes the results of the (albeit small) study even more compelling.

Enabling more expert-like behaviour

Medical information technologies should be designed to help the non-expert readers behave more like experts by enabling them to quickly find and focus their attention on the parts of the information important for their task. Our project has indicated that adding a Summary of Findings table to a systematic review may compensate for non-expert levels of numeracy in health professionals, helping them to access the main results in the review more quickly and comprehend them correctly.

CONCLUSION

The final version of the Summary of Findings table improved the understanding and rapid retrieval of the key findings of the review compared to reviews with no SoF table. However, the trial we conducted was small and generalisability of the results is uncertain. The Cochrane Handbook now recommends that review authors include a SoF table in their review[36]. These SoF tables are a "work in progress" and will continue to be developed and improved upon. Work is currently in progress on producing and evaluating versions of SoF tables for use in summaries targeted at consumers and policy makers. Future work includes development of SoF tables for diagnostic accuracy reviews.

ACKNOWLEDGMENTS

We thank Jan Ødegaard-Jensen for his great help with the statistical analysis and Arild Bjørndal for help with the manuscript.

.

REFERENCES