ELSEVIER

Journal of Clinical Epidemiology 63 (2010) 607−619

Journal of Clinical Epidemiology

ORIGINAL ARTICLES

User testing and stakeholder feedback contributed to the development of understandable and useful Summary of Findings tables for Cochrane reviews

Sarah E. Rosenbaum[a,*], Claire Glenton[b], Hilde Kari Nylund[a], Andrew D. Oxman[a]

[a]*Norwegian Knowledge Centre for the Health Services, PO Box 7004, St. Olavs plass, N-0130 Oslo, Norway*
[b]*Department of Global Health, Sintef, Oslo, Norway*

Accepted 22 December 2009

Abstract

Objective: To develop a Summary of Findings (SoF) table for use in Cochrane reviews that is understandable and useful for health professionals, acceptable to Cochrane Collaboration stakeholders, and feasible to implement.

Study Design and Setting: We gathered stakeholder feedback on the format and content of an SoF table from an advisory group of more than 50 participants and their constituencies through e-mail consultations. We conducted user tests using a think-aloud protocol method, collecting feedback from 21 health professionals and researchers in Norway and the UK. We analyzed the feedback, defined problem areas, and generated new solutions in brainstorming workshops.

Results: Stakeholders were concerned about precision in the data representation and about production feasibility. User testing revealed unexpected comprehension problems, mainly confusion about what the different numbers referred to (class reference). Resolving the tension between achieving table precision and table simplicity became the main focus of the working group.

Conclusion: User testing led to a table more useful and understandable for clinical audiences. We arrived at an SoF table that was acceptable to the stakeholders and in principle feasible to implement technically. Some challenges remain, including presenting continuous outcomes and technical/editorial implementation.

Keywords: Knowledge translation; Health numeracy; Risk communication; Systematic reviews; Information design; Usability; User experience; Fuzzy traces theory

Everything should be made as simple as possible but not simpler. (*Albert Einstein*)

Simplicity is highly overrated. (*Donald Normann*)

1. Background

Limited time is a frequently cited barrier to clinicians' use of evidence in practice [1−6]. Systematic reviews help to address this problem by summarizing evidence [7] but are still too time consuming to be practical for busy professionals. Further summarization of systematic reviews could help make evidence more useful and easy to find for decision makers with limited time [8,9].

This is the first of two articles on the development and evaluation of summaries of Cochrane reviews for clinicians and other typical users of The Cochrane Library or general medical journals [10]. The challenges and solutions we discuss here are also likely to be relevant for other systematic reviews and health technology assessments.

Summarized evidence for clinicians exists in many different formats, for instance as structured abstracts, synopses published in secondary journals, and online services. Haynes' 5-S pyramid describes a typology of increasingly condensed and clinically useful formats: from studies (and their abstracts) to syntheses (systematic reviews), synopses (eg, ACP Journal Club), summaries (eg, Clinical Evidence), and systems (eg, EPJ reminders) [11−14]. The PRISMA statement [15] provides a consensus-based checklist for producing abstracts for systematic reviews, and Hartley [16] reviews how these abstracts might be made clearer for a wide target group. However, little research has been published describing how clinicians experience summaries of systematic reviews. Numerical presentations of risk can be difficult, even for highly educated populations [17]. On the other

* Corresponding author. Tel.: +47-98256051; fax: +47-23255010.
E-mail address: sarah@rosenbaum.no (S.E. Rosenbaum).

0895-4356/$ − see front matter
doi: 10.1016/j.jclinepi.2009.12.013

What is new?

Key finding:

- We have developed a Summary of Findings (SoF) table for presenting results from systematic reviews that strikes a balance between precision and simplicity.

What this adds to what is known?

- How results are presented in SoF tables (including details about numerical representation and text and visual formatting) strongly influence users' perceptions and understanding of the data.

What are the implications, what should change now?

- Numbers in a table that need to be compared should belong to the same class. All numbers should be labeled explicitly so that class reference becomes apparent. Enabling easy gist extraction may also make the table less error prone.

hand, risk communication studies have shown that text-based descriptions of the effect of an intervention tend to be interpreted inconsistently by different people [18–20] and that numbers may be preferred by people making important health care decisions [21].

Earlier work on creating summaries of Cochrane reviews has also illustrated that to summarize already synthesized evidence is challenging and can lead to misrepresentation of the original data [22]. When attempting to summarize evidence for consumers with back pain, researchers encountered several critical issues:

- Large numbers of reported outcomes made it difficult to identify those outcomes that are most clinically relevant.
- Critical information was missing, for example, information about adverse effects and scales.
- Lack of standardization in the numerical presentation of results, the qualitative description of these results, and the manner in which the quality of data was evaluated made understanding difficult.

The GRADE system offers possible solutions to some of these challenges. GRADE is a structured, transparent system that allows authors to evaluate and report the quality of evidence [23,24]. An output of GRADE is a "Summary of Findings" (SoF) table, where authors are encouraged to focus on the most important outcomes, including those outcomes with no data or statistically nonsignificant data and adverse effects. Authors' judgments about the quality of evidence are presented together with the results for each outcome. The GRADE SoF table offers a useful starting point for summary authors by bringing the most important information to the foreground, regardless of the results or lack of them, and explicitly highlighting the quality of the evidence for each outcome.

Since 2004, open discussions have taken place in the Cochrane Collaboration about including SoF tables in Cochrane reviews [25], and extensive input has been gathered from stakeholders on the content and formatting of such tables. However, a number of issues continued to remain unresolved. A working group was, therefore, established to continue developing an SoF table designed for inclusion in Cochrane reviews and to evaluate this table.

The SoF table should summarize the key results of the review by presenting what is known and not known about the benefits and harms of an intervention, as well as how sure we can be of the evidence. It should be understandable and useful for a clinical audience, without oversimplifying or incorrectly presenting the data. We also needed to ensure that the content and data presentation was acceptable to Cochrane stakeholders and that the formatting was feasible to produce within the technical constraints of the system for publishing Cochrane reviews. In this article, we present and discuss the development process that led to our final decisions regarding table content, format, and data representation. In a second article [10], we present the effect of including a table in a Cochrane review on user satisfaction, understanding, and time spent finding key results.

2. Methods

To develop a table that works for different types of data, we searched for a Cochrane review that included dichotomous and continuous outcomes and outcomes with no data. The Cochrane review on the effect of compression stockings for preventing deep vein thrombosis in airline passengers [26] had all these types of results. It also covered a topic that was of potential interest to many people, making it easy to use in an evaluation process involving participants with different backgrounds. Using GRADE, we generated an SoF table for this review (Fig. 1).

We used cycles of multiple methods to develop the table:

- Advisory group feedback to inform table development from a stakeholder perspective.
- User testing methods to inform table development from a user perspective.
- Brainstorming workshops to generate ideas and solutions to problems uncovered through feedback and testing.

We also carried out two randomized controlled trials (RCTs) between development cycles to measure user satisfaction, correct understanding, and time spent to find main messages in the review, the results of which are reported in another article [10]. We fed all stakeholder and user feedback into the brainstorming workshops. For an overview of the entire process, see Fig. 2.