Characteristics of the Cochrane Oral Health Group Systematic Reviews

Sorin T. Teich, DMD, MBA; Lisa A. Lang, DMD, MBA; Catherine A. Demko, PhD

Abstract: The Cochrane Oral Health Group (COHG) was formed in 1994 with the aim of producing systematic reviews that primarily include oral health randomized controlled trials (RCTs). The purpose of this cross-sectional study was to characterize reviews published by the COHG. In September 2013, the COHG database was accessed, and all publications were downloaded. Reviews with no studies identified according to the inclusion criteria were labeled “empty reviews.” The complete Cochrane database included a total of 5,697 reviews, of which the COHG database included 142 reviews. Of these 142, 69 (48.6%) did not reach a conclusion, including 20 (14.1%) that were identified as empty reviews. Of the 122 non-empty reviews, 116 (95.1%) were based exclusively on RCTs. The median number of RCTs and patients included in the non-empty reviews were seven and 489, respectively. The median number of included RCTs and patients for reviews that reached conclusions were 12 and 934, respectively, and there were five RCTs and 211 patients for reviews without conclusions. Overall, the characteristics of the Cochrane oral health reviews (OH-CSRs) were similar to Cochrane reviews published in other disciplines (All-CSRs). The authors observed a significant difference in the median number of RCTs and patients included when reviews that reached conclusions were separated from those that did not. A greater proportion of empty reviews were present in OH-CSRs compared with All-CSRs. Turning the Cochrane reviews into a tool that is more relevant in clinical practice will require implementation of a methodology allowing inclusion of non-RCTs while controlling for possible bias.

Dr. Teich is Associate Dean of Clinical Operations and Associate Professor, Department of Comprehensive Care, Case Western Reserve University School of Dental Medicine; Dr. Lang is Assistant Dean of Clinical Education and Associate Professor, Department of Comprehensive Care, Case Western Reserve University School of Dental Medicine; and Dr. Demko is Associate Professor, Department of Community Dentistry, Case Western Reserve University School of Dental Medicine. Direct correspondence to Dr. Sorin T. Teich, Department of Comprehensive Care, Case Western Reserve University School of Dental Medicine, 2124 Cornell Rd., Cleveland, OH 44106; 216-368-6161; sorin@case.edu.

Keywords: dental research, evidence-based dentistry, evidence-based practice, Cochrane reviews, systematic reviews, randomized controlled trials, controlled clinical trials, empty reviews

Submitted for publication 5/2/14; accepted 5/30/14

The dental profession is still looking for the best approach to teach future dentists the concepts of evidence-based dentistry (EBD). The American Dental Education Association (ADEA)’s Competencies for the New General Dentist require that graduates should be competent to “evaluate and integrate best research outcomes with clinical expertise and patient values for evidence-based practice,” therefore encouraging integration of EBD principles into dental curricula. The Commission on Dental Accreditation (CODA)’s predoctoral standards state that “Graduates must be competent to access, critically appraise, apply, and communicate scientific and lay literature as it relates to providing evidence-based patient care” and that “Patient care must be evidence-based, integrating the best research evidence and patient values.”

In the 1990s, a process of critically reviewing the scientific literature for evidence emerged in medicine. A pyramid of evidence developed by Sackett et al. characterized systematic reviews as being the apex of the pyramid. This hierarchy was further refined by Glover et al. and Rosner who characterized the highest level of evidence as “filtered information.” In this category, the highest ranked studies are systematic reviews that gather data from the literature, preferably conducted in randomized controlled trials (RCTs).

The Cochrane Collaboration was formed in 1992 in response to Archie Cochrane’s call for practitioners to use evidence from RCTs because they are most likely to provide reliable information. The importance of the Cochrane Collaboration was compared to the Human Genome Project, with the collaboration said to have the potential to be the most significant clinical steppingstone since the founding of the National Institutes of Health (NIH) in the United States. The original task of the col-
laboration was to “prepare, maintain, and disseminate systematic, up-to-date reviews of RCTs of health care, and, when RCTs are not available, reviews of the most reliable evidence from other sources.”12 This statement is still aligned with the latest version of the Cochrane Handbook, which states the organization’s aim is “to help people make well-informed decisions about health care . . . by providing a reliable synthesis of the available evidence on a given topic, . . . and facilitate decisions considering all the evidence on the effect of an intervention” (emphasis added).13 Recently, the collaboration approved a new mission statement affirming that “our mission is to promote evidence-informed health decision-making by producing high-quality, relevant, accessible systematic reviews and other synthesized research evidence.”14 The organization set ten principles to guide its work towards fulfilling the mission, but it is interesting that the term “RCT” is no longer to be found in either the mission or the principles. Formed in 1994, the Cochrane Oral Health Group’s current aim is “to produce systematic reviews which primarily include all RCTs of oral health,”15 but this goal seems to be disconnected from both the Cochrane Collaboration’s mission and the organization’s aim as outlined in the handbook.

Comparisons of Cochrane systematic reviews (CSRs) with non-Cochrane systematic reviews (NCSRs) have been carried out in many disciplines. In the first years after the collaboration was founded, there was debate as to whether the quality of CSRs was as low or better than NCSRs.16-18 More recently, the consensus is that CSRs are generally of a higher quality because of better bias control,19,20 higher update rates,17,21 and controlled methodologies.20,22 Despite this evidence, physicians’ ratings of CSRs for relevance and newsworthiness have been reported to be significantly lower than NCSRs.23 Physicians also have been found to be less likely to access the full text of CSRs,24 a behavior hypothesized to pertain to their length and tendency to be less conclusive.24,25 The high number of reviews with no definite conclusions and recommendations for further research, together with a structure that contains irrelevant material and is not conducive to easy consultation for physicians seeking answers to clinical questions, are also considered contributing factors for undervaluation of CSRs.25 The amplitude of the problem is illustrated in a report published in 2007 noting that 96% of the CSRs recommended further research and 49% were inconclusive regarding the available evidence supporting certain interventions.26 Other predictors of utilization of best-available evidence in medicine and CSRs in particular are not directly related to their structure and include perceived intervention effects, clinical experience, and employment rank and environment.27

The dental profession faces similar challenges driven by the need to integrate with other medical disciplines and to provide care based on best available evidence. It has been shown that although some dental specialties are better overall at reporting research, all systematic reviews in dentistry need improvement.28,29 A study that evaluated dental practitioners’ attitudes about evidence-based practice (EBP) determined that only 29% of the dentists could correctly define EBP.30 When faced with clinical uncertainties, that study reported that 60% of general practitioners turned to colleagues and not to the literature to get help. This attitude towards EBP is also prevalent among specialists, with one study reporting that orthodontists “were most likely to change their practice philosophy based on expert advice.”31 That study also found that 59% of the participating orthodontists characterized the literature as ambiguous and 55% were not aware of the existence of the Cochrane Collaboration. This result is surprising because it has been shown that CSRs have the best quality and present the highest form of evidence in comparison with NCSRs in the orthodontic discipline.32

The growth in the number of systematic reviews in dentistry is dramatic. In 2004, a survey of systematic reviews in dentistry identified a total of 131 articles (CSRs and NCSRs combined).33 In 2012, a total of 126 CSRs and 1,062 NCSRs was reported.34 The characteristics of systematic reviews have also evolved over the years: in 2004, 39% of total systematic reviews included only RCTs, 17% of the reviews did not find enough evidence to answer their key question, and an additional 50% hedged answering the key question by citing the weak quality of the supporting evidence.33 In 2012, 93.3% of the CSRs included RCTs only, compared with 17.6% for the NCSRs;14 however, 17.5% of the oral health CSRs (OH-CSRs) had no studies included and were thus defined as “empty reviews,”35 which is approximately twice the proportion in general medicine (All-CSRs), reported to be between 8.7 and 9.3%.36,37

The purpose of our cross-sectional descriptive study was to characterize the OH-CSRs published through September 10, 2013. The project aimed to compare characteristics of these reviews with All-CSRs and identify differences in the characteristics of reviews with a definite conclusion and those without
not RCTs were identified, as well as those categorized as empty reviews. All empty reviews were analyzed for the total number of rejected studies and their characteristics.

**Methodology**

On September 10, 2013, two of the authors (STT and LAL) accessed the Cochrane Library Oral Health Group Database (http://ohg.cochrane.org/published-updates-reviews-and-protocols). All published reviews were downloaded, and the same two authors independently screened the abstract and full texts to document the following: 1) publication date; 2) type and number of included studies; 3) total number of patients included in each review (because only a limited number of reviews report this number, patient numbers reported in the “Characteristics of included studies” table in each review were summarized, and only patients who were included in the CSR analysis [i.e., finished the last period of follow-up] were tabulated); and 4) if the review reached a specific conclusion/recommendation based on the included material. Reviews that did not include patients (i.e., with no studies identified according to the inclusion criteria) were identified and labeled as empty reviews. The full text of the empty reviews was screened to identify studies that were excluded after full text review. Following this process, the same authors compared their findings; consensus was reached through additional joint review of specific items in cases in which discrepancies were observed.

The number of RCTs and the total number of patients included in each review were plotted in histograms. The number of RCTs in each review was plotted against the number of patients included in the review on a logarithmic scale, and the median point for both parameters was determined. In order to analyze the characteristics of reviews that did not reach conclusions versus those that did, the median numbers of RCTs and patients included (as well as the range for each parameter) were tabulated. Finally, the reviews that included research designs that were not RCTs were identified, as well as those categorized as empty reviews. All empty reviews were analyzed for the total number of rejected studies and their characteristics.

**Results**

As of September 2013, the Cochrane Database of Systematic Reviews included a total of 5,697 reviews,38 of which the OHG database included 142 oral-health related reviews (2.5%). Among these 142 reviews, 49 (34.5%) did not reach a conclusion but had studies included; an additional 20 OH-CSRs (14.1%) were identified as empty reviews.39-58 Four reviews59-62 reached a partial conclusion regarding one of the interventions discussed, whereas for other interventions no conclusion was presented. The remainder of 69 reviews reached a definite conclusion/recommendation.

Table 1 shows characteristics of all the reviews excluding the “empty reviews” and the four with partial conclusions. Overall, the median numbers of RCT studies and patients included in the reviews were seven (range 0-144) and 489 (range 14-191,873), respectively. Figure 1 and Figure 2 show the frequency and the cumulative percentages of the numbers of RCTs and patients included in the “non-empty” 122 reviews. On a logarithmic scale, Figure 3 shows the numbers of RCTs versus the number of patients included in these reviews.

When the reviews were separated based on presence/absence of a well-defined conclusion (N=69 and 49, respectively), it was evident that both the number of RCTs and the patients included in the reviews with conclusions were significantly larger than those reviews without conclusions. The median numbers of included RCTs and patients were 12 and 934, respectively, for reviews that reached conclusions, and five RCTs and 211 patients for reviews without conclusions (Table 1). Among the 122 non-empty reviews, 116 (95.08%) were based exclusively on RCTs, four also included other study designs, and two were based on observational studies.

| Table 1. Characteristics of all Cochrane oral health reviews through September 2013 excluding empty reviews |
|-------------------------------------------------|-------------------------------------------------|-------------------------------------------------|
| # of Reviews | Median # of RCTs Included (Range) | Median # of Patients Included (Range) |
| All reviews* | 122 | 7 (0-144) | 489 (14-191,873) |
| Reviews with conclusions | 69 | 12 (1-144) | 934 (23-191,873) |
| Reviews without conclusions | 49 | 5 (0-28) | 211 (20-2,302) |

*Four reviews with mixed conclusions were included only in the overall analysis.

RCT=randomized controlled trial.
Figure 1. Number of randomized controlled trials (RCTs) included in 122 Cochrane oral health systematic reviews (OH-CSRs)

Figure 2. Number of patients included in 122 Cochrane oral health systematic reviews (OH-CSRs)
types in addition to RCTs,\textsuperscript{63-66} and two were based exclusively on non-RCT study designs.\textsuperscript{67,68} The types of trials included in these last-mentioned six reviews are described in Table 2.

Table 3 shows the characteristics of OH-CSR empty reviews, including how many articles were rejected after reviewing the full text, the type of article that was rejected (clinical/non-clinical), and how many articles were rejected because the trial was not an RCT or the randomization method was not clear. Empty reviews are inconsistent in the way they report the rejected studies; therefore, no data were available for two reviews,\textsuperscript{42,51} and partial data were available for two others.\textsuperscript{47,48} One CSR did not review any study in full text.\textsuperscript{58}

**Discussion**

The results of this study are meaningful for the dental profession and dental education because treatments in dentistry (as in any other medical-related discipline) should be driven by best available evidence. Understanding the current status, limitations, and advantages of what is almost considered to be the “gold standard” of best available evidence summaries, i.e., the CSRs, will help overcome barriers\textsuperscript{30,31} to the evidence-based practice of dentistry.

Overall, 142 OH-CSRs published before September 2013 were mapped in this study. Since the total number of CSRs (oral and non-oral) published was 5,697, the oral-health related reviews comprised only 2.5\% of those. Considering the dental share of total health expenditure in the United States is around 3.7\%,\textsuperscript{69} one would have expected that the percentage of oral health-related reviews would be similar. This discrepancy was pointed out in a previous study in relation to medicine, which showed the economic burden of disease had only a modest correlation (\(r=0.54, p=0.014\)) with the number of systematic reviews in the Cochrane database.\textsuperscript{70}

Among the 142 oral health reviews, 122 contained studies appropriate for inclusion (Table 1).
### Table 2. Reviews that included other study designs beside RCTs

<table>
<thead>
<tr>
<th>Authors</th>
<th># of RCTs</th>
<th># of CCTs</th>
<th># of Cohort Studies</th>
<th># of Control Studies</th>
<th># of Cross-Sectional Surveys</th>
<th># of Patients</th>
<th>Conclusion=1, No conclusion=0</th>
</tr>
</thead>
<tbody>
<tr>
<td>Harrison and Ashby, 2001</td>
<td>5</td>
<td>7</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>462</td>
<td>1</td>
</tr>
<tr>
<td>Mandall et al., 2003</td>
<td>0</td>
<td>3</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>157</td>
<td>0</td>
</tr>
<tr>
<td>Millett et al., 2007</td>
<td>5</td>
<td>3</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>488</td>
<td>0</td>
</tr>
<tr>
<td>Zakrzewska et al., 2005</td>
<td>8</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>470</td>
<td>0</td>
</tr>
<tr>
<td>Wong et al., 2010</td>
<td>2</td>
<td>0</td>
<td>1</td>
<td>6</td>
<td>16</td>
<td>30,925</td>
<td>1</td>
</tr>
<tr>
<td>Oliver et al., 2008</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>349</td>
<td>0</td>
</tr>
</tbody>
</table>

RCT=randomized controlled trial; CCT=controlled clinical trial  
Note: See References for full citations of specified reviews.

### Table 3. Characteristics of empty reviews and studies rejected after full-text review

<table>
<thead>
<tr>
<th>Authors</th>
<th>Total # of Rejected Studies</th>
<th># of Rejected Clinical Studies</th>
<th># of Rejected In Vitro Studies</th>
<th># of Rejected Other Studies</th>
<th># of Clinical Studies Rejected Because Trial Is Non-RCT or Randomization Not Clear</th>
<th>Remarks</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pereira-Cenci et al., 2009</td>
<td>2</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>Del Fabbro et al., 2009</td>
<td>3</td>
<td>3</td>
<td>0</td>
<td>2</td>
<td>2</td>
<td></td>
</tr>
<tr>
<td>Pedrazzi et al., 2008</td>
<td>10</td>
<td>6</td>
<td>3</td>
<td>1 literature review</td>
<td>2</td>
<td></td>
</tr>
<tr>
<td>Parkin et al., 2008</td>
<td>6</td>
<td>6</td>
<td></td>
<td></td>
<td>2</td>
<td></td>
</tr>
<tr>
<td>Esposito and Worthington, 2013</td>
<td>14</td>
<td>14</td>
<td>NA</td>
<td></td>
<td>No table</td>
<td></td>
</tr>
<tr>
<td>Sharif, Catleugh, et al., 2010</td>
<td>16</td>
<td>16</td>
<td>10</td>
<td></td>
<td>10</td>
<td></td>
</tr>
<tr>
<td>Dorri et al., 2009</td>
<td>2</td>
<td>2</td>
<td></td>
<td></td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>Sharif et al., 2010</td>
<td>12</td>
<td>11</td>
<td>1 opinion letter</td>
<td></td>
<td>9</td>
<td></td>
</tr>
<tr>
<td>Eliyas et al., 2013</td>
<td>77</td>
<td>NA</td>
<td>77</td>
<td></td>
<td>No table; data from flow chart</td>
<td></td>
</tr>
<tr>
<td>Dashash et al., 2013</td>
<td>19</td>
<td>19</td>
<td></td>
<td></td>
<td>2</td>
<td></td>
</tr>
<tr>
<td>Yu et al., 2013</td>
<td>0</td>
<td>0</td>
<td></td>
<td></td>
<td>0 Flow chart shows no full-text evaluated</td>
<td></td>
</tr>
<tr>
<td>Millett et al., 2012</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td></td>
<td>No table or flow chart</td>
<td></td>
</tr>
<tr>
<td>Luther et al., 2010</td>
<td>4</td>
<td>4</td>
<td></td>
<td></td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>Sharif, Fedorowicz, et al., 2010</td>
<td>3</td>
<td>3</td>
<td></td>
<td></td>
<td>3</td>
<td></td>
</tr>
<tr>
<td>Sharif, Merry, et al., 2010</td>
<td>4</td>
<td>4</td>
<td></td>
<td></td>
<td>4</td>
<td></td>
</tr>
<tr>
<td>Innes et al., 2008</td>
<td>6</td>
<td>6</td>
<td></td>
<td></td>
<td>1</td>
<td></td>
</tr>
<tr>
<td>Belmonte et al., 2013</td>
<td>6</td>
<td>6</td>
<td></td>
<td></td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>Coulthard et al., 2010</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td></td>
<td>No table or flow chart</td>
<td></td>
</tr>
<tr>
<td>Ahangari et al., 2010</td>
<td>8</td>
<td>8</td>
<td></td>
<td></td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>de Souza et al., 2010</td>
<td>1</td>
<td>1</td>
<td></td>
<td></td>
<td>1</td>
<td></td>
</tr>
</tbody>
</table>

Note: See References for full citations of specified reviews. Items in Remarks column refer to table in the Cochrane review on “Characteristics of excluded studies” and flow chart describing search and inclusion/exclusion strategy.
The majority of these included only RCTs (95.1%). This is the same order of magnitude (98.4%) reported in relation to CSRs in all disciplines (All-CSR), but significantly different from the 17.6% reported for NCSRs.34,71

Fewer than five RCTs were included in 43.4% of the reviews (Figure 1), with the overall median number being seven. This number is close to that of other Cochrane specialty groups,36 in which the median number for All-CSRs is eight,26,71 but significantly lower than the median number of studies included in oral NCSRs, which is reported to be 15.34 It is interesting to note that 69 (48.6%) of the OH-CSRs reached a definite conclusion regarding interventions; this proportion closely agrees with that reported for All-CSRs (47.8%).26 However, when the median number of included RCTs in the OH-CSRs that reached conclusions was compared to the non-empty OH-CSRs without a conclusion, it became evident that the former included significantly more RCTs than did the latter (12 versus five) (Table 1).

A similar trend was observed when the number of patients included was summarized (Table 1). Overall, about 25% of the OH-CSRs included fewer than 100 patients (Figure 2), with a median number of 489. This number is lower than the reported median number of patients included in All-CSRs and in NCSRs (769 and 1,137 patients, respectively).71

Only six OH-CSRs also included other design studies (Table 2): four with RCTs63-66 and two without RCTs.67,68 Although only two reviews out of these six reached a definitive conclusion,63,65 these examples show that an OH-CSR is allowed to include studies that are not RCTs. This mirrors the disconnect between the aim of the Cochrane Oral Health Group, which mandates inclusion of RCTs,15 and the Cochrane Handbook, which mandates "providing a reliable synthesis of the available evidence on a given topic, . . . and facilitate decisions considering all the evidence on the effect of an intervention."13

Despite this statement, the vast majority of OH-CSRs and All-CSRs included exclusively RCTs; this led to about half of any CSR not reaching a definite conclusion regarding the intervention because not enough evidence was available. Furthermore, our review revealed that about 14% of the OH-CSRs were empty reviews, a slight reduction from previous reports,34,37 but still significantly higher than the proportion in All-CSRs (approximately 9%).36,37 It is interesting to note that the percentage of empty reviews in All-CSRs had nearly doubled since 2003 when it was only 5%.72

The high percentage of CSRs without a definitive conclusion can be interpreted as an illustration of Rossi’s “Stainless Steel Rule of Evaluation,” an empirical law stating that “the better designed the impact assessment . . . , the more likely is the resulting estimate of net impact to be zero”—i.e., more rigorous designs are more likely to show no effects.73 There has been much debate regarding whether RCTs and non-RCT studies provide results of similar quality. One study reported that the summary odds of prospective studies have good correlation with these obtained in the RCTs.74 “Well-designed” observational studies also did not systematically overestimate the magnitude of the effects of treatment.75,76 Another study, however, did not concur with these conclusions and supported the idea that randomization is essential to minimize bias and distortion of the outcome of care.77 Even when the question of randomization is not taken into account, the lack of standardization of criteria for levels of evidence may cause the results of a set of particular studies to vary when different criteria are applied.78 A previous report showed that CSRs are less prone to bias than NCSRs that perform meta-analyses, but whether “such bias is likely to affect the conclusions of a systematic review . . . must be carefully assessed for each case.”79

We should also consider that an RCT cannot be designed for every clinical question, as pointed out by Smith and Pell in their satirical but nonetheless important study.80 Economics, patient preferences and compliance, and ethical considerations may create a situation that will preclude conducting an RCT.81 Empty reviews are a particular case of systematic reviews that do not reach a definitive conclusion. Our results show (Table 3) that most empty OH-CSRs rejected an average of 10.7 (SD=17.4) clinical studies. This number corresponds to the average number of articles excluded from empty All-CSRs, i.e., 9.6 (SD=14.5).37

Excluding the 77 studies rejected in the empty review by Eliyas et al.47 (the authors do not report how many of these were clinical studies), we gather from the column “# of Rejected Clinical Studies” in Table 3 that 110 clinical trials were rejected after full text review. Of these, it is reported that 39 (35.5%) were rejected solely because the randomization method was not clear or the study was not designed as an RCT. We also observed a lack of standardization in the empty OH-CSRs in relation to reporting how the rejected articles were excluded. Only one82 out of 20 empty OH-CSRs (5%) reported no studies excluded; this is significantly lower than the 23.4% reported in the All-CSRs empty reviews.37
The main questions that arise are these: do the empty reviews have any value, and should these reviews even be published? Lang et al. present the opinion that empty CSRs should be published, but propose that the reviewers “offer some of their observations generated through the review . . . even if none of them meet eligibility criteria.” These authors also suggest considering if the inclusion criteria were too strict, especially for issues that “do not lend themselves to experimental design.” In response, Green et al. submit that this approach is not recommended because “drawing conclusions . . . from studies that do not meet the a priori inclusion criteria has the potential to increase the risk of bias.” It has also been suggested that “absence of studies at the RCT level should not paralyze decision making.” With nearly 85% of the Cochrane Review Groups hosting at least one empty review and the percentage of empty OH-CSRs being significantly higher than in other disciplines, it is evident that the reviewers in the Cochrane Collaboration think that these reviews have value despite having the potential to cause undervaluation of CSRs as an aid for clinical decision making.

The approach that postulates including only RCTs in the systematic review therefore results in ignoring the results of a potentially well-designed research project. Further, a side effect of economic inefficiency is created. In fiscal year 2010-11, the NIH spent approximately $3.4 billion or 11% of its overall budget on RCTs. The overall budget spent on applied research (non-basic biomedical research) was 45% of the total budget of $31 billion, i.e., nearly $14 billion. In choosing to include only RCTs, the Cochrane Collaboration may be excluding quality evidence generated in NIH-funded clinical trials worth billions of dollars per year. These numbers do not account for non-NIH-funded translational and clinical research that is probably even more likely to be designed as a non-RCT.

A solution for the dilemma of empty reviews is to return to the original intent of the Cochrane Collaboration. Since nothing precludes inclusion of studies that are not designed as RCTs in a CSR, these studies should at least be evaluated for quality and bias. Ironically, the major roadblock that precludes including non-RCTs is found in the same version of the Cochrane Handbook that advocates “considering all the evidence.” In 2008, the Cochrane Collaboration released the “Risk of Bias” (ROB) tool (Table 8.5.a in the handbook) that is designed to assess RCTs only. Although the ROB tool is the standard for Cochrane reviews, it proved to have low reliability for evaluating the risk of bias, grading of evidence, and interpretation of that evidence. Among the six OH-CSRs that included other studies beside RCTs (Table 2), only one was published after the ROB tool was implemented.

If the concern is that non-RCT clinical trials cannot be reported because the a priori inclusion criteria were not met, the solution is to design a clear, progressive, a priori methodology that allows inclusion of non-RCTs in cases in which analysis of the results of RCTs only is not conclusive or no RCTs are available. Clearly, potential bias risks induced by including non-RCTs should be evaluated and discussed. To address this need, a new tool for evaluating the risk of bias for nonrandomized studies (RoBANS) is available and has shown validity. The RoBANS tool was developed to be compatible with the Cochrane ROB and the hope is that it will be available to Cochrane reviewers, allowing them to include also non-RCTs in CSRs.

Conclusion

Overall, this study found that characteristics of the OH-CSRs were similar to All-CSRs when considering the median number of included RCTs, the proportion of reviews that reach a definite conclusion, and the proportion of reviews that include only RCTs. We observed a significant difference in the median number of RCTs and patients included when we separated reviews that reached conclusions from those that did not.

Empty reviews are an extreme case of reviews without conclusions, and a greater proportion of these reviews were present in OH-CSRs compared with All-CSRs. Although the specific reason for this finding is not clear, we suggest that the disconnection between the aims of the Cochrane Oral Health Group and the Cochrane Collaboration may be a contributing factor.

According to the collaboration’s mission and the aims noted in the handbook, its basic approach is to explore all best-available evidence; in reality, however, the vast majority of the reviews are limited to inclusion of RCTs. It is clear that this selectivity causes a significant proportion of CSRs to be nonconclusive. We believe that the economic inefficiencies we have pointed out in this article result from the collaboration’s decision to apply a study inclusion strategy that is disconnected from its own mission.
These inefficiencies are not sustainable and are at the root of previous reports that question the relevance and use of CSRs in medicine and dentistry.

Turning the Cochrane reviews into a tool that is more conclusive and relevant to clinicians, educators, and policy makers will require implementation of a stratified a priori methodology for inclusion of non-RCTs, at least in the cases in which a CSR is empty. This strategy also will incorporate adequate tools for bias estimation for non-RCTs, allowing reviewers and readers to weigh the implications of including studies with “lower” level methodologies. We believe that this direction will be embraced with enthusiasm by Cochrane’s customers: those who will read the reviews and subsequently apply the conclusions for the benefit of their patients.

REFERENCES


43. Dashash M, Yeung C, Jamous I, Blinkhorn A. Interventions for the restoration of teeth that have been weakened by the absence of enough covering of enamel, caused by amelogenesis imperfecta. Cochrane Database Syst Rev 2013(6).

44. de Souza RF, Travess H, Newton T, Marchesan MA. Interventions for treating traumatized ankylosed permanent front teeth. Cochrane Database Syst Rev 2010(1).


