

The effects of teaching strategies on learning to think critically in primary and secondary schools: protocol for an overview of systematic reviews

Oxman AD, Dahlgren A, García LM, Kaseje M, Nsangi A, Rosenbaum S, Semakula D, Sewankambo NK

Working paper, December 2019

Colophon

- Title* The effects of teaching strategies on learning to think critically in primary and secondary schools: protocol for an overview of systematic reviews
- Authors* Oxman, Andrew D
Dahlgren, Astrid
García, Laura Martínez
Kaseje, Margaret
Nsangi, Allen
Rosenbaum, Sarah
Semakula, Dan
Sewankambo NK
- Corresponding author(s)* Oxman, Andy
oxman@online.no
Centre for Informed Health Choices
Norwegian Institute of Public Health
Postboks 222 Skøyen
0213 Oslo, Norway
- Keywords* teaching instructional, pedagogical, strategies, techniques, methods, approaches, critical thinking, primary school, secondary school, systematic reviews
- Citation* Oxman AD, Dahlgren A, García LM, Kaseje M, Nsangi A, Rosenbaum S, Semakula D, Sewankambo NK. The effects of teaching strategies on learning to think critically in primary and secondary schools: protocol for an overview of systematic reviews. IHC Working Paper, 2019
- Article category* **Systematic reviews**
 Development and evaluation of learning resources
- Date* December 2019

Abstract

Background

An overview of what is known from systematic reviews of the effects of teaching strategies can inform the design of learning resources and decisions about which teaching strategies to use. We are undertaking this overview to inform decisions about which teaching strategies to use to teach critical thinking to primary and secondary school students.

Objectives

Our primary objective is to provide an overview of what is known from systematic reviews about the effects of strategies to help primary and secondary school students learn to think critically.

Methods

The overview will be conducted in two stages. In the first stage we will map characteristics of systematic reviews of teaching strategies. We will include reviews that assess the effects of teaching strategies that can potentially be used in primary or secondary schools to help students learn to think critically, have a “Methods” section with explicit selection criteria, report at least one outcome measure of the ability to undertake one of four basic types of cognitive tasks (memory, procedural, comprehension, or opinion), and were published within the past 20 years. In the second stage, we will prepare structured summaries of the systematic reviews that are most relevant to our primary objective and synthesize those findings.

Discussion

Teaching strategies may be difficult to evaluate; synthesising the findings of evaluations may be difficult; and the applicability of the findings of evaluations from one setting to another may be uncertain. Nonetheless, people making decisions about teaching strategies can be helped by an overview summarising what is known from available systematic reviews.

Background

Critical thinking

Learning to think critically is widely held to be an aim of education [1]. However, there is no agreement on the definition of “critical thinking”, or which frameworks (conceptual structures intended to serve as a support or guide) best support critical thinking [2-6]. Ennis has defined critical thinking as “reasoned, reflective thinking focused on deciding what to believe or do” [7], and we will use that definition.

Thinking evolved to help us choose what to do to achieve our goals after taking account of estimates of the likely effects of our actions [8]. A fundamental goal of critical thinking is to improve decision making by increasing the likelihood that we will believe and act on those claims that are more likely to help us achieve our goals [9].

Teaching strategies

Definitions of teaching (instructional, or pedagogical) strategies (techniques, methods, or approaches) vary. Some authors distinguish between strategies, techniques, methods, and approaches. However, there is overlap in how these terms are used. Our focus is on “different ways of helping students to learn - that is, different ways of helping them to achieve the learning outcomes that [teachers] have decided are important” [10]. We will refer to these as “teaching strategies”.

There are several lists of teaching strategies, organised in different ways. For example, Beck surveyed 25 teacher education textbooks and was unable to find two similar lists of teaching strategies [11]. Some examples of different teaching strategies are shown in Table 1. The (ERIC) thesaurus includes 117 terms under “teaching methods” and 44 terms under “instruction” (Appendix 1). On the other hand, the U.S. Institute of Education Sciences recommends only seven strategies to improve student learning [12]:

- Space learning over time
- Interleave worked example solutions and problem-solving exercises
- Combine graphics with verbal descriptions
- Connect and integrate abstract and concrete representations of concepts
- Use quizzing to promote learning

Table 1. Examples of teaching strategies

Beck 1998 taxonomy of teaching strategies [11]	Dorgu 2017 teaching methods [13]	Killen 2015 effective teaching strategies [10]	Examples of systematic reviews
<p>ASSOCIATIVE <i>Objective:</i> To group students based on their skills, needs and interests to help facilitate mental, emotional, and/or social growth <i>Examples:</i> dyads, partners, cross/multi-age, ability and interest groups, heterogeneous, homogeneous, cooperative learning, teams</p> <p>DELIBERATIVE <i>Objective:</i> To encourage a thoughtful exchange of ideas to promote cognitive, social, and verbal communication skills <i>Examples:</i> debate, round table, conference, panel, symposium, magic circle, fishbowl, brainstorm, buzz session, class discussion</p> <p>EXPOSITIVE <i>Objective:</i> To provide information, oral or written, in an orderly, authoritative, and intelligible manner, to a receptive audience <i>Examples:</i> lecture, recitation, review, oral or written report, textual readings, graphical materials, demonstration, modelling, testing</p> <p>INDIVIDUALISTIC <i>Objective:</i> To provide instruction designed to meet the skills, needs, and interests of the student, based on individual assistance <i>Examples:</i> programmed, self-paced, packet, contract, learning styles, mastery learning, independent study, tutorial, interest centres</p> <p>INTERROGATIVE <i>Objective:</i> To use questioning skills to encourage participation, clarify and evaluate understanding, and promote higher thinking <i>Examples:</i> convergent, divergent, prompt, probe, redirect, repetition, interview, open-ended, higher level, Socratic, questioning</p> <p>INVESTIGATIVE <i>Objective:</i> To solve problems, based on inductive reasoning, by collecting and analysing data, and drawing conclusions <i>Examples:</i> inquiry, exploration, problem solving, critical thinking, experimentation, laboratory, case study/method, discovery</p> <p>PERFORMATIVE <i>Objective:</i> To encourage creative, aesthetic, and/or psychomotor expression based on the dramatic/fine arts, and physical skills <i>Examples:</i> dramatic play, role play, storytelling, choral reading, calisthenics, dance, mock trial, rehearsal, simulation, gaming</p> <p>TECHNOLOGICAL <i>Objective:</i> To allow students to access and record information by means of mechanical devices, from film projectors to computers <i>Examples:</i> audio/videotaping, overhead/film projecting, televising, video conferencing, word processing, Internet, emailing</p>	<p>Cognitive development methods</p> <ul style="list-style-type: none"> • discussion method • questioning/Socratic method • team teaching method • talk chalk/recitation method • field trip/excursion method • team teaching method <p>Affective development methods</p> <ul style="list-style-type: none"> • modelling method • simulation method • dramatic method • simulation games • role-playing method <p>Psychomotor development methods</p> <ul style="list-style-type: none"> • inquiry method • discovery method • process approach method • demonstration method • laboratory/experimentation method • programmed learning method • Dalton plan/assignment method • project method • microteaching method • mastery learning 	<ul style="list-style-type: none"> • Direct instruction • Discussion • Small-group work • Co-operative learning • Problem solving • Inquiry • Role-play • Case study • Student writing <p>El Soufi 2019 teaching critical thinking [14]</p> <ul style="list-style-type: none"> • Instruction in general critical thinking skills (involves training students to define arguments, evaluate reliability of sources, identify fallacies and assumptions, use inductive and deductive logic, synthesise information, make inferences, etc. • Debate • Assessment techniques • Literary and narrative texts • Brainstorming techniques • Journal writing • Scaffolding • Active learning strategies <p>Caro 2016 (PISA) [15]</p> <ul style="list-style-type: none"> • Cognitive activation • Teacher directed strategies • Student oriented strategies <p>Alberta learning 2002 [16]</p> <p>Instructional strategies that are especially effective in the health education program include:</p> <ul style="list-style-type: none"> • cooperative learning • group discussion • independent study • portfolio development • journals and learning logs • role-playing • cognitive organizers • literature response • service learning • issue-based inquiry 	<ul style="list-style-type: none"> • Audio podcast [17] • Classroom discussion [18] • Clickers with different teaching strategies [19] • Collaborative learning [20,21] • Computer-based feedback [22] • Computer-based scaffolding [23] • Concept mapping-based learning technologies [24] • Cooperative learning [25] • Digital games [26] • Flipped classroom instructional strategy [27, 28] • Homework [29] • Humorous lectures [30] • Inquiry-based science teaching [31] • In-service teacher training interventions [32] • Problem- and project-based learning [33] • Serious games [34] • Small-group discussions in science teaching [35, 36] • Small group learning [37]

- Help students to allocate time efficiently
- Help students build explanations by asking and answering deep questions

Pomerance and her colleagues summarised these strategies slightly differently [38]:

- Pairing graphics with words
- Linking abstract concepts with concrete representations
- Posing probing questions
- Repeatedly alternating problems with their solutions provided and problems that students must solve
- Distributing practice
- Assessing to boost retention

They reviewed 48 textbooks for elementary and secondary teacher training and found that none of the textbooks accurately described those six fundamental instructional strategies. At most, only two of the six were covered in any textbook, and when textbooks did mention the strategies (allowing for a broad range of terminology and descriptions), the discussion could be as brief as 1-2 sentences.

Learners and learning outcome measures

Although our specific interest is in primary and secondary school students and critical thinking outcomes, we will not initially limit this overview to that population or those outcome measures. There are four reasons for this. First, there are not many reviews that focus specifically on critical thinking [39-50]. To the extent that those reviews do consider the effects of specific strategies, they tend to be broad categories and comparisons of strategies tend to be made indirectly (in meta-regression analyses). For example, Abrami and colleagues explored differences in the effect of three types of instruction (dialogue; authentic or anchored instruction; and mentoring, coaching, or tutoring) across 341 comparisons with different populations, outcome measures, and study designs [45]. Thus, an overview that only included critical thinking as an outcome would be of limited value.

Second, although some learning outcomes may be of little relevance to learning to think critically, it is difficult to specify a priori which outcomes are completely irrelevant and which might provide useful information despite not directly measuring critical thinking. For example, on the one hand it can be argued that outcome measures that only require retention of knowledge are irrelevant to critical thinking. On the other hand, it is important that students have knowledge of Key Concepts (principles for critical thinking) and that they retain that knowledge. Other outcome measures, such as reading comprehension or

understanding of science texts, are dependent on a range of factors in addition to critical thinking.

Third, many reviews are not limited to primary or secondary school interventions and may or may not explore differences in effects across different learners. Although some teaching strategies might be expected to have different effects for different types of learners, it is uncertain whether this is the case. Starting out with an overly narrow focus in terms of the learners could result in an overview that is far less informative than it might otherwise be.

Fourth, it is uncertain how many potentially useful systematic reviews of teaching strategies there are and what the characteristics of those reviews are.

For these reasons, we will first conduct a mapping overview, to characterise the range of systematic reviews of teaching strategies that can potentially inform the design of resources to help primary and secondary school students learn to think critically. The mapping overview can be useful to teachers and others with an interest in other learning outcomes and learners. It will also enable us to make an informed decision about which reviews are likely to be most useful for our specific interests and to focus on those.

Why it is important to do this overview

We want to provide an overview of what is known from systematic reviews about the effects of strategies that can be used to help primary and secondary school students learn to think critically. Our immediate aim is to inform the design of learning resources to teach lower secondary school students in East Africa to think critically about health claims and choices. This overview can also help teachers, teacher trainers, and other educators to identify effective strategies for teaching critical thinking. It will also help to identify needs and priorities for evaluations of teaching strategies, as well as priorities for systematic reviews of the effects of teaching strategies. The overview will also help to develop a framework for considering alternative teaching strategies.

Objectives

We are undertaking this overview to inform decisions about which teaching strategies to use to teach critical thinking to primary and secondary school students. The primary objective is to provide an overview of what is known from systematic reviews about the effects of strategies to help primary and secondary school students learn to think critically.

Secondary objectives are to:

- map the characteristics of systematic reviews of teaching strategies

- identify needs and priorities for evaluations of teaching strategies based on the findings of the included systematic reviews
- identify needs and priorities for systematic reviews of the effects of teaching strategies for which we are unable to find reliable, up-to-date systematic reviews
- inform the development of a framework for types of teaching strategies

Methods

Criteria for considering systematic reviews for inclusion

In the first stage of this overview we will map characteristics of systematic reviews of teaching strategies. We will include systematic reviews that:

- assess the effects of teaching strategies (different ways of helping students to learn) that can potentially be used in primary or secondary schools to help students learn to think critically,
- have a “Methods” section with explicit selection criteria,
- report at least one outcome measure of the ability to undertake one of four basic types of cognitive tasks (memory, procedural, comprehension, or opinion) [51], and
- were published within the past 20 years.

We will exclude reviews of teaching strategies that are restricted to:

- professional students (e.g. medical or nursing students) other than teacher training
- special education (teaching children and youth with disabilities)
- creative or physical skills such as artistic, cooking, musical or physical skills

Doyle [51] defined the four basic types of cognitive tasks noted above as follows:

1. memory tasks in which students are expected to recognize or reproduce information previously encountered (e.g., memorize a list of spelling words or lines from a poem);

2. procedural or routine tasks in which students are expected to apply a standardized and predictable formula or algorithm to generate answers (e.g., solve a set of subtraction problems);

3. comprehension or understanding tasks in which students are expected to (a) recognize transformed or paraphrased versions of information previously encountered, (b) apply procedures to new problems or decide from among several procedures those which are applicable to a particular problem (e.g., solve "word problems" in mathematics), or (c) draw inferences from previously encountered information or procedures (e.g., make predic-

tions about a chemical reaction or devise an alternative formula for squaring a number);

4. opinion tasks in which students are expected to state a preference for something (e.g., select a favorite short story).

These tasks roughly correspond with Blooms taxonomy, which has six main categories of intellectual abilities and skills [52]. Bloom's taxonomy is well known and has clear definitions, but it difficult to make clear distinctions between the higher-order categories [4]. For the purposes of this overview, we will consider any task that requires judgement ('evaluation' in Bloom's taxonomy) as 'opinion tasks' including judgements about what to believe and what to do.

Search methods for identification of systematic reviews

We have created an initial list of potentially relevant teaching strategies (Box 1) by reviewing several lists [10-16,36,53]. We started with Beck's taxonomy [11], which we have adapted and reorganised, considering other teaching strategies and ways of categorising these. We will continue to develop this list of terms iteratively, based on the literature that we retrieve and input from educational researchers and teachers.

Box 1. List of teaching strategies

Didactic strategies (instruction in which information is presented directly from the teacher to the student, in which the teacher selects the topic, controls instructional stimuli, obligates a response from the student, evaluates responses, and provides reinforcement for correct responses and feedback for incorrect ones) [54]

Direct instruction, lectures, textbooks, picture books, audio-visual aids, podcasts, multimedia instruction, demonstration, modelling, mini lessons, reading, graphic presentations, combined graphic and verbal presentations, narrative text, comics, humour, scaffolding, pre-teaching vocabulary, link abstract concepts with concrete representations

Questioning techniques (methods used for constructing and presenting questions in order to promote effective discussions and learning or to elicit information) [55-57]

Socratic method, open ended questions, closed questions, interviewing, prompting, probing, redirecting, wait time, clickers, pose probing questions, oral or written reports, cloze, "assess to boost retention", quizzes, ask and answer deep questions

Discussion strategies

Classroom discussion, small group discussion, buzz sessions, brainstorming, round table, debate, structured controversy, magic circle, fishbowl dialogue, four sides/corners strategy, reflective discussion, flipped classroom

Role playing

Read aloud, readers' theatre, dramatic play, storytelling, mock trial, simulation, learning games, public speaking and speech writing

Problem-based learning

Enquiry-based learning, exploration-based learning, student research, research projects, learning through experimentation, science fairs, science Olympics, using case studies to teach, laboratory

teaching methods, field trips, discovery learning, analytic memo, concept attainment, concept formation, concept maps, graphic organizer, knowledge map, cognitive organiser, mind mapping, structured overview, “repeatedly alternating problems with their solutions provided and problems that students must solve”

Repetition and progression

Distributed practice, space learning over time, spaced learning, pacing, learning targets, learning progression, competency-based learning, sequential approach, explicit teaching, interdisciplinary teaching

Assessment techniques [58-61]

Feedback, classroom assessment techniques, formative assessment, background knowledge probe, the one-minute paper, traffic light cards, muddiest point, what’s the principle, problem recognition task, student generated test questions, classroom opinion polls, directed paraphrasing, pro and con grid, student goals ranking, course-related interest and skills checklist, self-diagnostic learning logs, misconception/preconception check, empty outlines, invented dialogues, diagnostic teaching, precision teaching

Collaborative learning (a situation in which two or more people learn or attempt to learn something together)

[62]

Dyads, partners, cross/multi-age groups, ability and interest groups, heterogeneous groups, homogeneous groups, cooperative learning, heads together strategy, numbered heads together strategy, jigsaw teaching technique, team learning, peer teaching, peer partner learning, reciprocal teaching, readers’ workshop, reading buddies, think-ink-pair-share learning strategy, think-pair-share learning strategy, heterogeneous grouping, homogeneous grouping, multiple intelligences activities

Individual learning

individualized instruction, learner-controlled instruction, self-paced learning, independent study, programmed learning, contract learning, mastery learning, tutorial instruction, learning centres, menus, course packets, teaching tailored to students’ learning styles, Dalton plan, writing, writing to inform, paraphrasing, pause and reflect, journal writing, homework, practice, anchor activities

E-learning (using electronic devices, applications, or processes to acquire or transfer knowledge, attitudes, or skills through study, instruction, or experience) [55]

Online learning, web-based learning, web-based instruction, Web Quest, computer-based training, mobile learning, virtual classrooms, webinars, interactive e-lessons, online discussions, electronic simulations, audio/video recording

In-service teacher training

Microteaching, powerful pedagogical strategies, team teaching, scaffolding, peer teaching, teachers’ guides, or any of the other teaching strategies listed above

Starting with the terms in Box 1, we have developed search strategies for Education Research Complete (EBSCO) (Appendix 2) and for Education Resources Information Center (ERIC) (Ovid) (Appendix 3). In addition, we will search a database of systematic reviews collected by the Education Endowment Foundation (Steven Higgins, personal communication 7 October 2019).

Selection of systematic reviews

Two authors will independently screen the titles using Covidence [63] to identify systematic reviews that meet our inclusion criteria for the mapping overview.

Disagreements will be resolved by discussion, involving a third author if needed. We will pilot test the selection criteria on a sample of 100 records as training and to develop additional guidance, if needed, before screening the search results. We will retrieve the full text of articles that appear to meet the selection criteria and two authors will independently assess each article for inclusion.

Data collection

For each systematic review included in the mapping overview two authors will independently collect the following data:

- Types of students (primary school, secondary school, higher education, mixed, not clear)
 - If mixed, is it possible to extract effect estimates for primary or secondary school students from the reported results? (yes, no)
- Teaching strategies that were compared (using terminology used by the review authors)
- Cognitive tasks that were measured (memory, procedural, comprehension, opinion, mixed, not clear)
 - If 'opinion', whether this included judgements about what to believe or what to do (critical thinking)
- Other outcome measures that were reported
- Types of included studies (randomised trials, non-randomised studies, mixed)
- Date of the last search for studies
- Assessment of risk of bias (done using explicit criteria, considered without explicit criteria, not done)

Based on these data, we will make a judgement regarding the relevance of the review to our primary objective (directly relevant and informative, likely relevant and informative, possibly relevant and informative, probably not relevant, not relevant). These judgements will be discussed by the review team and a consensus will be reached on the systematic reviews that are most relevant to our primary objective.

One author will assess the reliability of each of those reviews using criteria developed by the SUPPORT and Supporting the Use of Research Evidence (SURE) collaborations (Appendix 4). Based on these criteria each review will be categorised as having:

- Only minor limitations
- Limitations that are important enough that it would be worthwhile to search for another systematic review and to interpret the results of this review cautiously, if a better review cannot be found
- Limitations that are important enough to make the findings of the review unreliable and so should be excluded from the overview

These judgements will be checked independently by a second author. Disagreements will be resolved by discussion, involving a third author if needed. We will then summarise each included systematic review using an approach developed by the SUPPORT Collaboration [64]. We will adapt template and guidelines for SUPPORT Summaries (Appendices 5 and 6) to prepare the summaries. The certainty of the evidence for the main comparisons will be assessed using the GRADE approach [65, 66] (Appendix 7). Each completed summary will be peer-reviewed (Appendix 8), checked by an editor (Appendix 9), and published on an open access website (e.g. www.informedhealthchoices.org).

Data synthesis

For each included systematic review, we will prepare a table summarising what the review authors searched for and what they found. We will prepare summary of findings tables for each main comparison and we will assess the relevance of the findings for teaching critical thinking in primary and secondary schools. The summaries will include key messages, important background information, a summary of the findings of the review and structured assessments of the relevance of the review. For this overview, our assessments of relevance will focus on the applicability of the findings to teaching critical thinking in primary and secondary schools, modifying factors (effect modifiers), resource use, and implications for practice and research. The summaries will be sent to the lead author of each review, at least one educational researcher, and at least one teacher with practical experience. The authors of the summaries will respond to each comment and make appropriate revisions, and the summaries will be copy edited. An editor will determine whether the comments have been adequately addressed and the summary is ready for publication on the Informed Health Choices website (www.informedhealthchoices.org).

The systematic reviews will be organised using the framework in Box 1. This framework will be adjusted iteratively to ensure that all the included systematic reviews are appropriately categorised and that all relevant teaching strategies are included and organised logically. We will prepare a table listing the included systematic reviews and types of teaching strategies for which we have not been able to identify a reliable review published in the past 20 years. We will also prepare a table of excluded reviews. This will include systematic reviews that address a question for which another (more up-to-date or reliable) review was included, reviews that are more than 20 years old, and reviews with limitations judged sufficient to compromise the reliability of the review.

We will describe the characteristics of the included systematic reviews in a table recording the date of the last search, any important limitations, what the review authors searched for and what they found. Our detailed assessments of the

reliability of the included systematic reviews will be summarised in a separate table showing whether each criterion in Appendix 4 was met for each review. Our structured synthesis of the findings of our overview will be based on two tables. The main findings of each review will be summarised in a table that includes the key messages from each summary. In a second table we will report the direction of the effects and the certainty of the evidence for each of the following types of outcomes: critical thinking (as defined by the review authors, if reported), cognitive tasks (memory, procedural, comprehension, and opinion) other learner outcomes (e.g. competences, dispositions, behaviours), resource use, effect modifiers, teacher outcomes, adverse effects (not captured by undesirable effects on any of the preceding types of outcomes), and any other important outcomes (that do not fit into any of the preceding types of outcomes). The direction of results will be categorised as: a desirable effect, little or no effect, an uncertain effect (very low certainty evidence), no included studies, an undesirable effect, not reported (i.e. not specified as a type of outcome that was considered by the review authors), or not relevant (i.e. no plausible mechanism by which the teaching strategy could affect the type of outcome).

We will consider factors besides the findings of the included systematic reviews when drawing conclusions about implications for practice [67]. These include considerations related to the applicability of the findings and the feasibility of the teaching strategies. Our conclusions about implications for systematic reviews will be based on types of teaching strategies for which we were unable to find a reliable, up-to-date review and limitations of the systematic reviews included. Our conclusions about implications for future research will be based on the findings of the included systematic reviews [68].

Discussion

Teaching strategies may be difficult to evaluate; synthesising the findings of evaluations may be difficult; and the applicability of the findings of evaluations from one setting to another may be uncertain. However, well-designed evaluations are preferable to poorly designed evaluations; systematic reviews are preferable to unsystematic reviews; and using the findings of systematic reviews to inform decisions is preferable to using unsystematic reviews.

Other types of information, including context-specific information, and judgements, including judgements about the applicability of the findings of systematic reviews in a specific context, are needed in addition. Nonetheless, people making decisions about teaching strategies should be helped by an overview summarising the findings of available systematic reviews, by identifying important uncertainties found by those systematic reviews, and by identifying where new or updated systematic reviews are needed. This overview should also help to inform judgements about the relevance of the available evidence in a specific context.

Acknowledgements

We would like to thank the following people for feedback on an earlier version of this protocol and advice:

- Iain Chalmers, Associate Member, Centre for Informed Health Choices, Oslo, Norway
- Steven Higgins, Professor, School of Education, Durham University, UK
- Monica Melby-Lervåg, Professor, Faculty of Educational Sciences, University of Oslo, Norway
- Matt Oxman, Researcher, Norwegian Institute of Public Health, Oslo, Norway

We are grateful to Marit Johansen for developing the search strategies in Appendices 2 and 3.

References

1. Voogt J, Roblin NP. A comparative analysis of international frameworks for 21st century competences: implications for national curriculum policies. *J Curric Stud* 2012; 357:299-321.
2. Presseisen BZ. *Critical Thinking and Thinking Skills: State of the Art Definitions and Practice in Public Schools*. Philadelphia: Research for Better Schools, Inc., 1986.
3. Follman J. Critical thinking definitions. *Inquiry* 1991; 8:4-5.
4. Moseley D, Baumfield V, Elliott, et al. *Frameworks for Thinking: A Handbook for Teaching and Learning*. Cambridge: Cambridge University Press 2005.
5. Moore T. Critical thinking: seven definitions in search of a concept. *Studies in Higher Education* 2013; 38:506-22.
6. Gyenes A. Definitions of critical thinking in context. *Annals of Educational Studies* 2015; 20:17-25.
7. Ennis RH. Critical thinking: reflection and perspective: part I. *Inquiry* 2011; 26:4-18.
8. Sloman S, Fernbach P. *The Knowledge Illusion: Why We Never Think Alone*. New York: Penguin, 2018.
9. Baron J. *Thinking and Deciding*, 4th ed. New York: Cambridge University Press, 2008.
10. Killen R. *Effective Teaching Strategies: Lessons from Research and Practice*. 7th ed. Melbourne: Cengage, 2015.
11. Beck CR. A taxonomy for identifying, classifying, and interrelating teaching strategies. *J Gen Educ* 1998; 47:37-62.
12. Pashler H, Bain P, Bottge B, et al. *Organizing Instruction and Study to Improve Student Learning (NCER 2007-2004)*. Washington, DC: National Center for Education Research, Institute of Education Sciences, U.S. Department of Education, 2007.
13. Dorgu TE. Different teaching methods: a panacea for effective curriculum implementation in the classroom. *Int J Secondary Educ* 2015; 3:77-87.

14. El Soufi N, See BH. Does explicit teaching of critical thinking improve critical thinking skills of English language learners in higher education? A critical review of causal evidence. *Studies Educ Eval* 2019; 60:140-62.
15. Caro DH, Lenkeit J, Kyriakides L. Teaching strategies and differential effectiveness across learning contexts: evidence from PISA 2012. *Studies Educ Eval* 2016; 49:30-41.
16. Alberta Learning. Instructional strategies. In: *Health and Life Skills Guide to Implementation (K-9)*. Alberta: Alberta Learning, 2002.
<https://education.alberta.ca/media/482311/is.pdf>
17. Hew KF. Use of audio podcast in K-12 and higher education: a review of research topics and methodologies. *Educ Tech Res Dev* 2009; 57:333-57.
18. Murphy PK, Wilkinson IAG, Soter A, Hennessey MN, Alexander JF. Examining the effects of classroom discussion on students' comprehension of text: a meta-analysis. *J Educ Psychol* 2009; 101:740-64.
19. Liu C, Chen S, Chi C, et al. The effects of clickers with different teaching strategies. *J Educ Comp Res* 2017; 55:603-28.
20. Chen J, Wang M, Kirschner PA, Tsai CC. The role of collaboration, computer use, learning environments, and supporting strategies in CSCL: a meta-analysis. *Rev Educ Res* 2018; 88:799-843.
21. Kumar R. The effect of collaborative learning on enhancing student achievement: a meta-analysis. MA Thesis, Concordia University, Montreal, Canada, 2017.
22. van der Kleij FM, Feskens RCW, Eggen TJHM. Effects of feedback in a computer-based learning environment on students' learning outcomes: a meta-analysis. *Rev Educ Res* 2015; 85:475-511.
23. Kim NJ, Belland BR, Walker AE. Effectiveness of computer-based scaffolding in the context of problem-based learning for STEM education: Bayesian meta-analysis. *Educ Psychol Rev* 2018; 30:397-429.
24. Stevenson MP, Hartmeyer R, Bentsen P. Systematically reviewing the potential of concept mapping technologies to promote self-regulated learning in primary and secondary science education. *Educ Res Rev* 2017; 21:1-16.
25. Johnson DW, Johnson RT, Stanne MB. *Cooperative learning methods: a meta-analysis*. University of Minnesota, Minneapolis, MN, 2000.
26. Clark DB, Tammer-Smith EE, Killingsworth SS. Digital games, design, and learning: a systematic review and meta-analysis. *Rev Educ Res* 2016; 86:79-122.

27. Cheng L, Ritzhaupt AD, Antonenko P. Effects of the flipped classroom instructional strategy on students' learning outcomes: a meta-analysis. *Educ Tech Res Dev* 2019; 67:793–824.
28. van Alten DCD, Phielix C, Jansen J, Kester L. Effects of flipping the classroom on learning outcomes: a meta-analysis. *Educ Res Rev* 2019; <https://doi.org/10.1016/j.edurev.2019.05.003>.
29. Fan H, Xu J, Cai Z, He J, Fan X. Homework and students' achievement in math and science: A 30-year meta-analysis. *Educ Res Rev* 2017; 20:35e54.
30. Martin DM, Preiss RW, Gayle BM, Allen M. A meta-analytic assessment of the effect of humorous lectures on learning. In Gayle B, Preiss R, Burnell N, Allen M (eds.), *Classroom Communication and Instructional Processes: Advances Through Meta-analysis*. Mahwah, NJ: Lawrence Erlbaum Associates, 2006; 295-313.
31. Furtak EM, Seidel T, Iverson H, Briggs DC. Experimental and quasi-experimental studies of inquiry-based science teaching: a meta-analysis. *Rev Educ Res* 2012; 82:300-29.
32. Popova A, Evans DK, Arancibia V. *Inside in-service teacher training: what works and how do we measure it?* Washington, DC: Rise Programme, 2016.
33. Jensen KJ. A Meta-analysis of the effects of problem- and project-based learning on academic achievement in grades 6-12 populations. *Education Dissertations*. 7.
34. Wouters P, van Nimwegen C, van Oostendorp H. A meta-analysis of the cognitive and motivational effects of serious games. *J Educ Psychol* 2013; 105:249-65.
35. Bennett J, Lubben F, Hogarth S, Campbell B. A systematic review of the use of small-group discussions in science teaching with students aged 11-18, and their effects on students' understanding in science or attitude to science. Review summary. University of York, UK, 2004.
36. Hogarth S, Bennett J, Campbell B, Lubben F, Robinson A. A systematic review of the use of small-group discussions in science teaching with students aged 11-18, and the effect of different stimuli (print materials, practical work, ICT, video/film) on students' understanding of evidence: Review summary. University of York, UK, 2005.
37. Lou Y, Abrami PC, d-Apollonia S. Small group and individual learning with technology: a meta-analysis. *Rev Educ Res* 2001; 71:449-521.
38. Pomerance L, Greenberg J, Walsh K. *Learning about Learning: What Every New Teacher Needs to Know*. Washington, DC: National Council on Teacher Quality, 2016.

39. Ay SC, Orhan A. The effect of different critical thinking teaching approaches on critical thinking skills: a meta-analysis study. Pamukkale Üniversitesi Eğitim Fakültesi Dergisi, doi: 10.9779/pauefd.561742.
40. Lorencová H, Jarošová E, Avgitidou S, Dimitriadou C. Critical thinking practices in teacher education programmes: a systematic review. Stud High Educ 2019; 44:844-59.
41. Puig B, Blanco-Anaya PB, Bargiela IM et al. A systematic review on critical thinking intervention studies in higher education across professional fields. Stud High Educ 2019; 44:860-9.
42. Ngudgratoke S. A meta-analysis of relationships between teaching methods and critical thinking. Re-Thinking Teacher Professional Education: Using Research Findings for Better Learning: Yearbook of Teacher Education, ICET 2017.
43. DiPasquale J, Hunter WJ. Critical thinking in asynchronous online discussions: a systematic review. Can J Learn Technol 2018; 44:2.
44. Ismail NS, Harun J, Salleh SM, et al. Supporting students' critical thinking with a mobile learning environment: a meta-analysis. Proceedings of INTED2016 Conference 7th-9th March 2016, Valencia, Spain.
45. Abrami PC, Bernard RM, Borokhovski E, et al. Strategies for teaching students to think critically a meta-analysis. Rev Educ Res 2015; 85: 275–314.
46. Karabulut US. Identifying instructional methods of teaching critical thinking: A systematic review and analysis of three decades of literature. Georgia Soc Stud J 2015; 5:96-107.
47. Tiruneh DT, Verburch A, Elen J. Effectiveness of critical thinking instruction in higher education: a systematic review of intervention studies. High Educ Stud 2014; 4:1.
48. Niu L, Behar-Horenstein LS, Garvan CW. Do instructional interventions influence college students' critical thinking skills? A meta-analysis. Educ Res Rev 2013; 9:114-28.
49. ten Dam G, Volman M. Critical thinking as a citizenship competence: teaching strategies. Learn Instr 2004; 14:359-79.
50. Bangert-Drowns RL, Bankert E. Meta-analysis of effects of explicit instruction for critical thinking. Paper presented at the Annual Meeting of the American Educational Research Association (Boston, MA, April 16-20, 1990). <https://eric.ed.gov/?id=ED328614>
51. Doyle W. Academic work. Rev Educ Res 1983; 53:159-99.
52. Bloom BS (ed.). Taxonomy of Educational Objectives: The Classification of Educational Goals. Handbook 1: Cognitive Domain. New York: McKay, 1956.

53. Branco CW. 100 Instructional Strategies. Union, NJ: Kean University – College of Education, 2016.
<https://www.kean.edu/~tpc/Instructional%20Strategies.pdf>
54. Austin S. Didactic Approaches. In: Volkmar F.R. (eds) Encyclopedia of Autism Spectrum Disorders. New York, NY: Springer, 2013.
https://doi.org/10.1007/978-1-4419-1698-3_2005
55. Institute of Education Sciences. Education Resources Information Center (ERIC) Thesaurus. Washington, DC: Institute of Education Sciences, 2019.
<https://eric.ed.gov/>
56. Illinois Center for Innovation in Teaching & Learning. Questioning strategies. In: Teaching & Learning Resources. <https://citl.illinois.edu/citl-101/teaching-learning/resources/teaching-strategies/questioning-strategies>
57. Gilbert SW. Systematic Questioning: Taxonomies that develop critical thinking skills. *Sci Teach* 1992; 59:41-6.
58. Wikipedia contributors. Classroom Assessment Techniques. Wikipedia, The Free Encyclopedia. July 16, 2018, 02:37 UTC.
https://en.wikipedia.org/w/index.php?title=Classroom_Assessment_Techniques&oldid=850471183.
59. Mihram D. Classroom assessment techniques. University of California, Center for Excellence in Teaching, 2010.
http://www.crlt.umich.edu/sites/default/files/resource_files/ClassroomAssessmentTechniquesHopkins.pdf
60. Cross PK, Angelo TA. Classroom Assessment Techniques: A Handbook for Faculty. Board of Regents of The University of Michigan, 1988.
<http://files.eric.ed.gov/fulltext/ED317097.pdf>
61. Gupta P. Popular K-12 classroom assessment techniques teachers need to know. *Ed Tech Review* 2017; <https://edtechreview.in/trends-insights/insights/2868-classroom-assessment-techniques-for-teachers>.
62. Dillenbourg P. Collaborative Learning: Cognitive and Computational Approaches. *Advances in Learning and Instruction Series*. New York, NY: Elsevier Science, 1999.
63. Covidence systematic review software, Veritas Health Innovation, Melbourne, Australia. Available at www.covidence.org
64. Rosenbaum SE, Glenton C, Wiysonge CS, et al. Evidence summaries tailored to health policymakers in low- and middle-income countries. *Bull World Health Organ* 2011; 89:54-61.
65. Guyatt GH, Oxman AD, Kunz R, HJ, et al. What is 'quality of evidence' and why is it important to clinicians? *BMJ* 2008; 336:995-98.

66. Cochrane Effective Practice and Organisation of Care (EPOC). EPOC worksheets for preparing a Summary of Findings (SoF) table using GRADE. EPOC Resources for review authors, 2017.
<https://epoc.cochrane.org/resources/epoc-resources-review-authors>
67. Cochrane Effective Practice and Organisation of Care (EPOC). Implications for research. EPOC Resources for review authors, 2017.
<https://epoc.cochrane.org/resources/epoc-resources-review-authors>
68. Alonso-Coello P, Schünemann HJ, Moberg J, et al. GRADE Evidence to Decision (EtD) frameworks: a systematic and transparent approach to making well-informed healthcare choices. 1. Introduction. BMJ 2016; 353:i2016.

Appendices

1. Key words for teaching methods and instruction in the Education Resources Information Center (ERIC) Thesaurus
2. Education Research Complete, Ebsco search strategy
3. ERIC search strategy
4. SUPPORT Summaries checklist for making judgements about how much confidence to place in a systematic review
5. Guidelines for preparing SUPPORT Summaries
6. Worksheets for preparing a Summary of Findings (SoF) table using GRADE
7. SUPPORT Summary peer review form
8. SUPPORT Summary peer review form
9. SUPPORT Summary template

Appendix 1. Key words for teaching methods and instruction in the Education Resources Information Center (ERIC) Thesaurus

Teaching Methods (117 terms)

Autoinstructional Methods (1966 1980)
Autoinstructional Programs (1966 1980)
Branching (Programmed Instruction)
Instructional Methods
Integrated Teaching Method
Negative Practice (2004)
Orff Method (2004)
Orff Schulwerk Approach (2004)
Presentation Methods
Programed Instruction (1966 1994)
Programed Units (1966 1980)
Programmed Instruction
Programmed Learning
Programmed Self Instruction
Project Methods
Project Training Methods (1968 1980)
Rudolf Steiner Schools
Socratic Method
Steiner Education
Suzuki Method (2004)
Teaching Methodology
Teaching Practices
Teaching Procedures (1966 1980)
Teaching Systems
Teaching Techniques (1966 1980)
Waldorf Education
Waldorf Educational Method
Waldorf Method (2004)
Waldorf Schools
[Broader Terms]
Educational Methods
[Narrower Terms]
Audiolingual Methods
Blended Learning
Case Method (Teaching Technique)
Clinical Teaching (Health Professions)
Community Based Instruction (Disabilities)
Conventional Instruction
Creative Teaching
Cross Age Teaching
Demonstrations (Educational)
Diagnostic Teaching
Direct Instruction
Discussion (Teaching Technique)
Drills (Practice)
Experimental Teaching
Grammar Translation Method
Individualized Instruction
Kinesthetic Methods
Language Experience Approach
Learner Controlled Instruction
Lecture Method
Montessori Method
Multimedia Instruction
Oral Communication Method

Peer Teaching
Precision Teaching
Reciprocal Teaching
Reggio Emilia Approach
Scaffolding (Teaching Technique)
Sight Method
Suggestopedia
Team Teaching
Telephone Instruction
Thematic Approach
Training Methods
Web Based Instruction
Whole Language Approach
[Related Terms]
Advance Organizers
Class Organization
Classroom Techniques
Cloze Procedure
Coaching (Performance)
Computer Simulation
Concept Mapping
Contingency Management
Course Organization
Curriculum Implementation
Developmentally Appropriate Practices
Dramatic Play
Duplication
Educational Strategies
Individual Instruction
Instruction
Instructional Effectiveness
Instructional Films
Instructional Leadership
Integrated Activities
Intermode Differences
Laboratory Procedures
Large Group Instruction
Learning Modalities
Learning Modules
Learning Strategies
Looping (Teachers)
Mass Instruction
Methods Research
Pacing
Pedagogical Content Knowledge
Praxis
Prompting
Questioning Techniques
Reinforcement
Repetition
Role Playing
Science Course Improvement Projects
Sequential Approach
Simulation
Small Group Instruction
Teacher Characteristics

Teaching Guides
Teaching Machines
Teaching Models
Theory Practice Relationship
Tutorial Programs
Writing Across the Curriculum

Instruction (44 terms)

Audiovisual Instruction
Clothing Instruction
College Instruction
"Community Based Instruction (Disabilities)"
Computer Assisted Instruction
Computer Managed Instruction
Conventional Instruction
Cooking Instruction
Direct Instruction
English Instruction
Ethical Instruction
Field Instruction
Foods Instruction
Geography Instruction
Group Instruction
History Instruction
Home Instruction
Humanities Instruction
Individual Instruction
Individualized Instruction
Instruction
"Language of Instruction"
Large Group Instruction
Learner Controlled Instruction
Library Instruction
Mass Instruction
Mathematics Instruction
Multimedia Instruction
Native Language Instruction
Nutrition Instruction
Pronunciation Instruction
Reading Instruction
Remedial Instruction
"Research and Instruction Units"
Science Instruction
Second Language Instruction
Sewing Instruction
Small Group Instruction
Speech Instruction
Spelling Instruction
Telephone Instruction
Textiles Instruction
Web Based Instruction
Writing Instruction

Appendix 2. Education Research Complete, Ebsco search strategy

#	Query	Results
S1	DE ("teaching method" or "teaching methods" or "teaching methodology" or "teaching methodologies" or "teaching model" or "teaching models" or "teaching strategy" or "teaching strategies" or "teaching technique" or "teaching techniques" or "teaching approach" or "teaching approaches") or SU ("teaching method" or "teaching methods" or "teaching methodology" or "teaching methodologies" or "teaching model" or "teaching models" or "teaching strategy" or "teaching strategies" or "teaching technique" or "teaching techniques" or "teaching approach" or "teaching approaches") or KW ("teaching method" or "teaching methods" or "teaching methodology" or "teaching methodologies" or "teaching model" or "teaching models" or "teaching strategy" or "teaching strategies" or "teaching technique" or "teaching techniques" or "teaching approach" or "teaching approaches") or TI ("teaching method" or "teaching methods" or "teaching methodology" or "teaching methodologies" or "teaching model" or "teaching models" or "teaching strategy" or "teaching strategies" or "teaching technique" or "teaching techniques" or "teaching approach" or "teaching approaches") or AB ("teaching method" or "teaching methods" or "teaching methodology" or "teaching methodologies" or "teaching model" or "teaching models" or "teaching strategy" or "teaching strategies" or "teaching technique" or "teaching techniques" or "teaching approach" or "teaching approaches")	53,177
S2	DE ("instructional method" or "instructional methods" or "instructional methodology" or "instructional methodologies" or "instructional model" or "instructional models" or "instructional strategy" or "instructional strategies" or "instructional technique" or "instructional techniques" or "instructional approach" or "instructional approaches") or SU ("instructional method" or "instructional methods" or "instructional methodology" or "instructional methodologies" or "instructional model" or "instructional models" or "instructional strategy" or "instructional strategies" or "instructional technique" or "instructional techniques" or "instructional approach" or "instructional approaches") or KW ("instructional method" or "instructional methods" or "instructional methodology" or "instructional methodologies" or "instructional model" or "instructional models" or "instructional strategy" or "instructional strategies" or "instructional technique" or "instructional techniques" or "instructional approach" or "instructional approaches") or TI ("instructional method" or "instructional methods" or "instruction-	8,435

	al methodology" or "instructional methodologies" or "instructional model" or "instructional models" or "instructional strategy" or "instructional strategies" or "instructional technique" or "instructional techniques" or "instructional approach" or "instructional approaches") or AB ("instructional method" or "instructional methods" or "instructional methodology" or "instructional methodologies" or "instructional model" or "instructional models" or "instructional strategy" or "instructional strategies" or "instructional technique" or "instructional techniques" or "instructional approach" or "instructional approaches")	
S3	DE ("instruction method" or "instruction methods" or "instruction methodology" or "instruction methodologies" or "instruction model" or "instruction models" or "instruction strategy" or "instruction strategies" or "instruction technique" or "instruction techniques" or "instruction approach" or "instruction approaches") or SU ("instruction method" or "instruction methods" or "instruction methodology" or "instruction methodologies" or "instruction model" or "instruction models" or "instruction strategy" or "instruction strategies" or "instruction technique" or "instruction techniques" or "instruction approach" or "instruction approaches") or KW ("instruction method" or "instruction methods" or "instruction methodology" or "instruction methodologies" or "instruction model" or "instruction models" or "instruction strategy" or "instruction strategies" or "instruction technique" or "instruction techniques" or "instruction approach" or "instruction approaches") or TI ("instruction method" or "instruction methods" or "instruction methodology" or "instruction methodologies" or "instruction model" or "instruction models" or "instruction strategy" or "instruction strategies" or "instruction technique" or "instruction techniques" or "instruction approach" or "instruction approaches") or AB ("instruction method" or "instruction methods" or "instruction methodology" or "instruction methodologies" or "instruction model" or "instruction models" or "instruction strategy" or "instruction strategies" or "instruction technique" or "instruction techniques" or "instruction approach" or "instruction approaches")	1,128
S4	DE ("pedagogical method" or "pedagogical methods" or "pedagogical methodology" or "pedagogical methodologies" or "pedagogical model" or "pedagogical models" or "pedagogical strategy" or "pedagogical strategies" or "pedagogical technique" or "pedagogical techniques" or "pedagogical approach" or "pedagogical approaches") or SU ("pedagogical method" or "pedagogical methods" or "pedagogical methodology"	4,690

	<p>or "pedagogical methodologies" or "pedagogical model" or "pedagogical models" or "pedagogical strategy" or "pedagogical strategies" or "pedagogical technique" or "pedagogical techniques" or "pedagogical approach" or "pedagogical approaches") or KW ("pedagogical method" or "pedagogical methods" or "pedagogical methodology" or "pedagogical methodologies" or "pedagogical model" or "pedagogical models" or "pedagogical strategy" or "pedagogical strategies" or "pedagogical technique" or "pedagogical techniques" or "pedagogical approach" or "pedagogical approaches") or TI ("pedagogical method" or "pedagogical methods" or "pedagogical methodology" or "pedagogical methodologies" or "pedagogical model" or "pedagogical models" or "pedagogical strategy" or "pedagogical strategies" or "pedagogical technique" or "pedagogical techniques" or "pedagogical approach" or "pedagogical approaches") or AB ("pedagogical method" or "pedagogical methods" or "pedagogical methodology" or "pedagogical methodologies" or "pedagogical model" or "pedagogical models" or "pedagogical strategy" or "pedagogical strategies" or "pedagogical technique" or "pedagogical techniques" or "pedagogical approach" or "pedagogical approaches")</p>	
S5	<p>S1 OR S2 OR S3 OR S4</p> <p>[GENERAL terms in Descriptor (DE), Subject (SU), Author keywords (KW), Title (TI), Abstract (AB)]</p>	64,036
S6	<p>DE ("DIDACTIC method (Teaching method)" or "DIRECT instruction" or "SCAFFOLDING (Teaching method)" or "LECTURE method in teaching" or "WIT & humor in education" or "AUDIOVISUAL aids in education" or "AUDIOVISUAL aids in early childhood education" or "AUDIOVISUAL aids in elementary education" or "AUDIOVISUAL aids in secondary education" or "FILMSTRIPS in education" or TEXTBOOKS or "TEACHING aids" or "COMIC books, strips, etc., in education" or "PICTURES in education" or "MULTIMEDIA systems in education" or "MEDIA programs (Education)" or "AUDIOVISUAL education" or "TEACHING demonstrations" or "PASSIVE learning" or "CONVENTIONAL instruction")</p>	35,715
S7	<p>DE (QUESTIONING or "CLOSED questions" or "OPEN-ended questions" or "INQUIRY method (Teaching)" or "INQUIRY-based learning" or "SOCRATIC method" or "EDUCATIONAL tests & measurements" or "STUDENT response systems" or "PROMPTING (Education)" or "CLOZE procedure")</p>	23,446
S8	<p>DE ("DISCUSSION in education" or DIALOGIC teaching" or BRAINSTORMING or "FLIPPED classrooms")</p>	1,700

S9	DE ("ROLE playing" or "DRAMA in education" or "SIMULATED environment (Teaching method)" or "SIMULATION methods in education" or "SIMULATION games in education" or GAMES or "EDUCATIONAL games" or "VIDEO games in education" or "EDUCATIONAL toys" or "STORYTELLING in education" or STORYTELLING or "ANCHORED instruction" or "PUBLIC speaking" or "TEACHING games for understanding" or "EXPRESSIVE language")	17,191
S10	DE ("INQUIRY-based learning" or "PROBLEM-based learning" or "ACTIVE learning" or "PROBLEM solving" or "ANCHORED instruction" or "CASE method (Teaching)" or "CRITICAL thinking" or "VIGNETTES (Teaching technique)" or "SCIENCE fairs" or "PROJECT method in teaching" or "SCIENCE projects" or "SCHOOL field trips" or "STUDENT research" or EXPERIMENTS or "SCIENTIFIC experimentation" or "CLASSROOM activities" or "LEARNING by discovery" or "DISCOVERY methods" or "CONCEPT learning" or "CONCEPT mapping" or MNEMONICS)	54,031
S11	DE ("DRILLS (Education)" or "PACING strategies (Education)") or "REPETITION (Learning process)" or "PRACTICE effects" or "INTERDISCIPLINARY teams in education" or "EDUCATION benchmarking" or "LEARNING goals" or "SEQUENTIAL approach (Teaching method)" or "OUTCOME-based education" or "FLASH cards" or "EXPLICIT instruction")	7,626
S12	DE ("ASSESSMENT for learning (Teaching model)" or "FORMATIVE tests" or "FEEDBACK (Psychology)" or "DIAGNOSTIC teaching" or "PRECISION teaching" or "COMMON fallacies")	6,493
S13	DE ("COLLABORATIVE learning" or "LEARNING by teaching" or "PEER teaching" or "STUDENT teachers" or "MIXED age grouping (Education)" or "CROSS-age teaching" or "ABILITY grouping (Education)" or "MIXED ability grouping (Education)" or "WITHIN-class grouping (Education)" or "GROUP work in education" or "REGGIO Emilia approach (Early childhood education)" or "RECIPROCAL teaching" or "PAIRED reading")	21,961
S14	DE ("INDIVIDUALIZED instruction" or "STUDENT-centered learning" or "CHILD-centered education" or "PACING strategies (Education)" or "PROGRAMMED instruction" or "MASTERY learning" or "TUTORS & tutoring" or "CLASSROOM learning centers" or "DALTON laboratory plan" or HOMEWORK or "ANCHORED instruction" or "JOURNAL writing" or "OPEN learning")	15,171
S15	DE ("VIRTUAL schools" or CYBERSCHOOLS or "VIRTUAL reality in education" or "VIRTUAL classrooms" or "ELECTRONIC classrooms" or "ONLINE education" or TELE-	84,592

	COURSES or "COMPUTERS in education" or "MOBILE learning" or "MOBILE apps in education" or "CD-ROMs in education" or "TABLET computers in education" or "COMPUTERS in elementary education" or "COMPUTER assisted instruction" or "EDUCATIONAL technology" or WEBQUESTS or "VIDEO games in education" or "INTERNET in education" or "VIRTUAL field trips" or "INTERACTIVE learning" or "ELECTRONIC discussion groups" or "ONLINE chat" or "SOUND recordings in education" or "AUDIOTAPES in education")	
S16	DE ("TEACHER training" or MICROTEACHING or MICROCOUNSELING or "TEACHER training courses" or "TEACHING teams" or "SCAFFOLDING (Teaching method)" or "PEER teaching" or "TEACHING guides")	26,664
S17	S6 OR S7 OR S8 OR S9 OR S10 OR S11 OR S12 OR S13 OR S14 OR S15 OR S16 [Specific descriptor terms (DE)]	254,661
S18	TI ("classroom technique*" or "class room technique*" or ((didactic W0 (strateg* or method* or technique* or approach* or instruction*)) or ((direct or active) W0 (instruction* or teaching)) or scaffolding or "scaffolded instruction*" or (lecture W0 (strateg* or method* or technique* or approach*)) or ((passive or conventional or traditional) W0 (teaching or instruction* or learning)) or (teaching W0 (aid or aids or demonstration*)) or audiovisual or "audio visual" or humor or textbook or textbooks or "text book" or "text books" or comic or comics or picture or pictures or film of films or multimedia or media or podcast* or questioning or ((closed or "open ended") W0 question*) or "inquiry based" or (socratic W0 (strateg* or method* or technique* or approach*)) or "educational test*" or quiz or quizzes or "response system*" or clicker or clickers or prompting or prompts or cloze or discussion* or dialogue* or "dialogic teaching" or debate or debates or debating or brainstorming or "buzz session*" or "structured controversy" or "magic circle" or ((flipped or inverted) W0 classroom*) or "flip teaching" or "role play*" or drama or "dramatic play" or simulation or "educational game*" or "training game*" or "instructive game*" or "video game*" or "learning game*" or storytelling or "story telling" or "educational toy" or "educational toys" or "public speaking" or "speech writing" or "anchored instruction*" or "readers' theatre" or "mock trial" or "mock trials" or (("problem based" or "inquiry based" or "enquiry based" or "exploration based" or "project based" or active) W0 (education or learning or teaching or training)) or (("problem based" or "inquiry based" or "enquiry based" or "exploration based" or "project based") W0 (technique* or method* or ap-	946,130

<p>proach* or strateg*) or (problem NO solving) or "case method*" or (critical NO (think* or reflect*)) or vignette* or "science fair*" or "science olympics" or "field trip*" or "project method*" or "project based method*" or "analytic memo*" or mnemonics or "memory training" or "memorization technique*" or "classroom activit*" or "learning by discovery" or "discovery learning" or (concept W0 (learning or mapping)) or drills or pacing or repetition* or "sequential approach*" or "flash card*" or "distributed practice" or "space learning" or "spaced learning" or "learning progression" or "interdisciplinary team*" or "inter disciplinary team*" or "interdisciplinary teaching" or "inter disciplinary teaching" or benchmarking or benchmark* or "learning goal*" or (("outcome based" or "competency based" or "results based" or "performance based" or "skills based") W0 (education or learning or teaching or training)) or ((explicit or detailed) W0 (instruction* or teaching)) or ((diagnostic or prescriptive or precision) W0 teaching) or "assessment technique*" or "formative assessment*" or "formative test*" or feedback or "feed back" or (common W0 (fallacy" or fallacies" or error* or blunder* or mistake* or misconception* or preconception*)) or "directed paraphrasing" or "traffic light card*" or ((collaborative or cooperative or "co operative" or team) W0 learning") or "peer teaching" or "peer teacher*" or "student teacher*" or (("mixed age" or "cross age") W0 teaching) or (("mixed age" or "cross age" or ability or "within class" or heterogeneous or homogenous) W0 group*) or "group work*" or "reciprocal teaching" or "jigsaw teaching" or "paired reading" or ((individual* or differentiated or personalized or programmed or anchored) W0 instruction*) or "anchor activity" or "anchored activities" or (("child controlled" or "student controlled" or "learner controlled") W0 (teaching or education or learning or instruction*)) or (("child centered" or "student centered" or "learner centered" or "self paced" or selfpaced or tailored) W0 (education or teaching or training or learning or instruction*)) or ((individual or independent or mastery or open or classroom or contract) W0 learning) or "dalton plan" or "dalton laboratory plan" or homework or "home work" or (journal* NO writing) or "course packet*" or "e learning" or "virtual school*" or cyberschool* or "cyber school*" or "virtual classroom*" or "electronic classroom*" or "virtual reality" or ((online or "on line" or "web based" or "computer based" or mobile or interactive) W0 (education or teaching or training or learning or instruction*)) or "educational technolog*" or internet or apps or "cdrom or "cd rom*" or computer* or "video game*" or webinar* or webquest* or "web quest*" or tele-</p>

	course* or "tele course*" or "electronic lesson*" or "e lesson*" or "electronic discussion*" or "online chat*" or "on line chat*" or "sound recording" or "audio recording" or "video recording" or audiotap* or (teacher NO training) or microteaching or "micro teaching" or microcounseling or "micro counseling" or "peer teaching" or "teaching team*" or "team teaching" or "teaching guides" or scaffolding or scaffolded)	
S19	AB ("classroom technique*" or "class room technique*" or ((didactic W0 (strateg* or method* or technique* or approach* or instruction*)) or ((direct or active) W0 (instruction* or teaching)) or scaffolding or "scaffolded instruction*" or (lecture W0 (strateg* or method* or technique* or approach*)) or ((passive or conventional or traditional) W0 (teaching or instruction* or learning)) or (teaching W0 (aid or aids or demonstration*)) or audiovisual or "audio visual" or humor or textbook or textbooks or "text book" or "text books" or comic or comics or picture or pictures or film of films or multimedia or media or podcast* or questioning or ((closed or "open ended") W0 question*) or "inquiry based" or (socratic W0 (strateg* or method* or technique* or approach*)) or "educational test*" or quiz or quizzes or "response system*" or clicker or clickers or prompting or prompts or cloze or discussion* or dialogue* or "dialogic teaching" or debate or debates or debating or brainstorming or "buzz session*" or "structured controversy" or "magic circle" or ((flipped or inverted) W0 classroom*) or "flip teaching" or "role play*" or drama or "dramatic play" or simulation or "educational game*" or "training game*" or "instructive game*" or "video game*" or "learning game*" or storytelling or "story telling" or "educational toy" or "educational toys" or "public speaking" or "speech writing" or "anchored instruction*" or "readers' theatre" or "mock trial" or "mock trials" or (("problem based" or "inquiry based" or "enquiry based" or "exploration based" or "project based" or active) W0 (education or learning or teaching or training)) or (("problem based" or "inquiry based" or "enquiry based" or "exploration based" or "project based") W0 (technique* or method* or approach* or strateg*)) or (problem NO solving) or "case method*" or (critical NO (think* or reflect*)) or vignette* or "science fair*" or "science olympics" or "field trip*" or "project method*" or "project based method*" or "analytic memo*" or mnemonics or "memory training" or "memorization technique*" or "classroom activity*" or "learning by discovery" or "discovery learning" or (concept W0 (learning or mapping)) or drills or pacing or repetition* or "sequential approach*" or "flash card*" or "distributed practice" or "space learning" or "spaced learning" or "learning pro-	946,210

	<p>gression" or "interdisciplinary team*" or "inter disciplinary team*" or "interdisciplinary teaching" or "inter disciplinary teaching" or benchmarking or benchmark* or "learning goal*" or (("outcome based" or "competency based" or "results based" or "performance based" or "skills based") W0 (education or learning or teaching or training)) or ((explicit or detailed) W0 (instruction* or teaching)) or ((diagnostic or prescriptive or precision) W0 teaching) or "assessment technique*" or "formative assessment*" or "formative test*" or feedback or "feed back" or (common W0 (fallacy" or fallacies" or error* or blunder* or mistake* or misconception* or preconception*)) or "directed paraphrasing" or "traffic light card*" or ((collaborative or cooperative or "co operative" or team) W0 learning)) or "peer teaching" or "peer teacher*" or "student teacher*" or (("mixed age" or "cross age") W0 teaching) or (("mixed age" or "cross age" or ability or "within class" or heterogeneous or homogenous) W0 group*) or "group work*" or "reciprocal teaching" or "jigsaw teaching" or "paired reading" or ((individual* or differentiated or personalized or programmed or anchored) W0 instruction*) or "anchor activity" or "anchored activities" or (("child controlled" or "student controlled" or "learner controlled") W0 (teaching or education or learning or instruction*)) or (("child centered" or "student centered" or "learner centered" or "self paced" or selfpaced or tailored) W0 (education or teaching or training or learning or instruction*)) or ((individual or independent or mastery or open or classroom or contract) W0 learning) or "dalton plan" or "dalton laboratory plan" or homework or "home work" or (journal* N0 writing) or "course packet*" or "e learning" or "virtual school*" or cyberschool* or "cyber school*" or "virtual classroom*" or "electronic classroom*" or "virtual reality" or ((online or "on line" or "web based" or "computer based" or mobile or interactive) W0 (education or teaching or training or learning or instruction*)) or "educational technolog*" or internet or apps or "cdrom or "cd rom*" or computer* or "video game*" or webinar* or webquest* or "web quest*" or telecourse* or "tele course*" or "electronic lesson*" or "e lesson*" or "electronic discussion*" or "online chat*" or "on line chat*" or "sound recording" or "audio recording" or "video recording" or audiotap* or (teacher N0 training) or microteaching or "micro teaching" or microcounseling or "micro counseling" or "peer teaching" or "teaching team*" or "team teaching" or "teaching guides" or scaffolding or scaffolded)</p>	
S20	KW ("classroom technique*" or "class room technique*" or ((didactic W0 (strateg* or method* or technique* or approach* or instruction*)) or ((direct or active) W0 (in-	946,127

<p>struction* or teaching)) or scaffolding or "scaffolded instruction*" or (lecture W0 (strateg* or method* or technique* or approach*)) or ((passive or conventional or traditional) W0 (teaching or instruction* or learning)) or (teaching W0 (aid or aids or demonstration*)) or audiovisual or "audio visual" or humor or textbook or textbooks or "text book" or "text books" or comic or comics or picture or pictures or film of films or multimedia or media or podcast* or questioning or ((closed or "open ended") W0 question*) or "inquiry based" or (socratic W0 (strateg* or method* or technique* or approach*)) or "educational test*" or quiz or quizzes or "response system*" or clicker or clickers or prompting or prompts or cloze or discussion* or dialogue* or "dialogic teaching" or debate or debates or debating or brainstorming or "buzz session*" or "structured controversy" or "magic circle" or ((flipped or inverted) W0 classroom*) or "flip teaching" or "role play*" or drama or "dramatic play" or simulation or "educational game*" or "training game*" or "instructive game*" or "video game*" or "learning game*" or storytelling or "story telling" or "educational toy" or "educational toys" or "public speaking" or "speech writing" or "anchored instruction*" or "readers' theatre" or "mock trial" or "mock trials" or (("problem based" or "inquiry based" or "enquiry based" or "exploration based" or "project based" or active) W0 (education or learning or teaching or training)) or (("problem based" or "inquiry based" or "enquiry based" or "exploration based" or "project based") W0 (technique* or method* or approach* or strateg*)) or (problem N0 solving) or "case method*" or (critical N0 (think* or reflect*)) or vignette* or "science fair*" or "science olympics" or "field trip*" or "project method*" or "project based method*" or "analytic memo*" or mnemonics or "memory training" or "memorization technique*" or "classroom activit*" or "learning by discovery" or "discovery learning" or (concept W0 (learning or mapping)) or drills or pacing or repetition* or "sequential approach*" or "flash card*" or "distributed practice" or "space learning" or "spaced learning" or "learning progression" or "interdisciplinary team*" or "inter disciplinary team*" or "interdisciplinary teaching" or "inter disciplinary teaching" or benchmarking or benchmark* or "learning goal*" or (("outcome based" or "competency based" or "results based" or "performance based" or "skills based") W0 (education or learning or teaching or training)) or ((explicit or detailed) W0 (instruction* or teaching)) or ((diagnostic or prescriptive or precision) W0 teaching) or "assessment technique*" or "formative assessment*" or "formative test*" or feedback or "feed back" or (common W0 (fallacy"</p>
--

	<p>or fallacies" or error* or blunder* or mistake* or misconception* or preconception*)) or "directed paraphrasing" or "traffic light card*" or ((collaborative or cooperative or "co operative" or team) W0 learning")) or "peer teaching" or "peer teacher*" or "stu- dent teacher*" or (("mixed age" or "cross age") W0 teaching) or (("mixed age" or "cross age" or ability or "within class" or heterogeneous or homogenous) W0 group*) or "group work*" or "reciprocal teaching" or "jigsaw teaching" or "paired reading" or ((individual* or differentiated or personalized or programmed or anchored) W0 in- struction*) or "anchor activity" or "anchored activities" or (("child controlled" or "stu- dent controlled" or "learner controlled") W0 (teaching or education or learning or in- struction*)) or (("child centered" or "student centered" or "learner centered" or "self paced" or selfpaced or tailored) W0 (education or teaching or training or learning or instruction*)) or ((individual or independent or mastery or open or classroom or con- tract) W0 learning) or "dalton plan" or "dalton laboratory plan" or homework or "home work" or (journal* N0 writing) or "course packet*" or "e learning" or "virtual school*" or cyberschool* or "cyber school*" or "virtual classroom*" or "electronic classroom*" or "virtual reality" or ((online or "on line" or "web based" or "computer based" or mobile or interactive) W0 (education or teaching or training or learning or instruction*)) or "educational technolog*" or internet or apps or "cdrom or "cd rom*" or computer* or "video game*" or webinar* or webquest* or "web quest*" or tele- course* or "tele course*" or "electronic lesson*" or "e lesson*" or "electronic discus- sion*" or "online chat*" or "on line chat*" or "sound recording" or "audio recording" or "video recording" or audiotap* or (teacher N0 training) or microteaching or "micro teaching" or microcounseling or "micro counseling" or "peer teaching" or "teaching team*" or "team teaching" or "teaching guides" or scaffolding or scaffolded)</p>	
S21	S18 OR S19 OR S20 [Specific terms in Title (TI), Abstract (AB), Author keywords (KW)]	946,218
S22	S5 OR S17 OR S21	1,011,777
S23	(DE ("systematic review*" or "meta-analysis" or metaanalysis)) OR (SU ("systematic review*" or "meta-analysis" or metaanalysis)) OR (KW ("systematic review*" or "me- ta-analysis" or metaanalysis)) OR (TI ("systematic review*" or "meta-analysis" or metaanalysis)) OR (AB ("systematic review*" or "meta-analysis" or metaanalysis))	12,531
S24	S22 AND S23	4,011

Appendix 3. ERIC search strategy

#	Searches	Results
1	active learning/ or associative learning/ or blended learning/ or cartoons/ or "case method (teaching technique)"/ or classroom techniques/ or cloze procedure/ or computer assisted instruction/ or computer games/ or computer uses in education/ or concept mapping/ or concept teaching/ or conventional instruction/ or cooperative learning/ or counseling techniques/ or creative teaching/ or critical thinking/ or cross age teaching/ or "demonstrations (educational)"/ or diagnostic teaching/ or direct instruction/ or discovery learning/ or "discussion (teaching technique)"/ or distance education/ or dramatic play/ or "drills (practice)"/ or educational games/ or educational methods/ or educational strategies/ or educational technology/ or electronic classrooms/ or electronic learning/ or experimental learning/ or experimental teaching/ or "feedback (response)"/ or field experience programs/ or field instruction/ or field trips/ or game based learning/ or games/ or group instruction/ or "grouping (instructional purposes)"/ or homework/ or independent study/ or individual instruction/ or individualized education programs/ or individualized instruction/ or inquiry/ or instruction/ or instructional films/ or instructional materials/ or integrated activities/ or large group instruction/ or journal writing/ or learner controlled instruction/ or learning activities/ or learning strategies/ or learning/ or lecture method/ or mass instruction/ or mastery learning/ or misconceptions/ or mixed age grouping/ or montessori method/ or multimedia instruction/ or online courses/ or pacing/ or peer teaching/ or problem based learning/ or problem solving/ or prompting/ or puzzles/ or questioning techniques/ or reciprocal teaching/ or Reggio Emilia approach/ or reinforcement/ or repetition/ or role playing/ or rote learning/ or "scaffolding (teaching technique)"/ or sequential approach/ or sequential learning/ or simulation/ or small group instruction/ or student centered learning/ or student journals/ or student participation/ or student projects/ or teaching guides/ or teaching methods/ or teaching models/ or team teaching/ or thematic approach/ or training methods/ or video games/ or virtual classrooms/ or web based instruction/ [ERIC index term]	526861
2	(teaching method or teaching methods or teaching methodology or teaching methodologies or teaching model or teaching models or teaching strategy or teaching strategies or teaching technique or teaching techniques or teaching approach or teaching approaches).tw.	201146
3	(instructional method or instructional methods or instructional methodology or instructional methodologies or instructional model or instructional models or instructional strategy or instructional strategies or instructional technique or instructional techniques or instructional approach or instructional approaches).tw.	17880

4	(instruction method or instruction methods or instruction methodology or instruction methodologies or instruction model or instruction models or instruction strategy or instruction strategies or instruction technique or instruction techniques or instruction approach or instruction approaches).tw.	1535
5	(pedagogical method or pedagogical methods or pedagogical methodology or pedagogical methodologies or pedagogical model or pedagogical models or pedagogical strategy or pedagogical strategies or pedagogical technique or pedagogical techniques or pedagogical approach or pedagogical approaches).tw.	4690
6	(classroom technique* or class room technique* or scaffolding or scaffolded instruction* or audiovisual or audio visual or humor or textbook or textbooks or text book or text books or comic or comics or picture or pictures or film of films or multimedia or media or podcast* or questioning or inquiry based or educational test* or quiz or quizzes or response system* or clicker or clickers or prompting or prompts or cloze or discussion* or dialogue* or dialogic teaching or debate or debates or debating or brainstorming or buzz session* or structured controversy or magic circle or flip teaching or role play* or drama or dramatic play or simulation or educational game* or training game* or instructive game* or video game* or learning game* or storytelling or story telling or educational toy or educational toys or public speaking or speech writing or anchored instruction* or readers theatre or mock trial or mock trials or case method* or vignette* or science fair* or science olympics or field trip* or project method* or project based method* or analytic memo* or mnemonics or memory training or memorization technique* or classroom activit* or learning by discovery or discovery learning or drills or pacing or repetition* or sequential approach* or flash card* or distributed practice or space learning or spaced learning or learning progression or interdisciplinary team* or inter disciplinary team* or interdisciplinary teaching or inter disciplinary teaching or benchmarking or benchmark* or learning goal* or assessment technique* or formative assessment* or formative test* or feedback or feed back or directed paraphrasing or traffic light card* or peer teaching or peer teacher* or student teacher* or group work* or reciprocal teaching or jigsaw teaching or paired reading or anchor activity or anchored activities or dalton plan or dalton laboratory plan or homework or home work or course packet* or e- learning or virtual school* or cyberschool* or cyber school* or virtual classroom* or electronic classroom* or virtual reality or educational technolog* or internet or apps or cdrom or cd rom* or computer* or video game* or webinar* or webquest* or web quest* or telecourse* or tele course* or electronic lesson* or e-lesson* or electronic discussion* or online chat* or on line chat* or sound recording or audio recording or video recording or audiotap* or microteaching or micro teaching or microcounseling or micro counseling or peer teaching or teaching team* or team teaching or teaching guides or scaffolding or	529898

	scaffolded).tw.	
7	(didactic adj (strateg* or method* or technique* or approach* or instruction*)).tw.	434
8	((direct or active) adj (instruction* or teaching)).tw.	3444
9	(lecture adj (strateg* or method* or technique* or approach*)).tw. [tw=abstract, title, heading word, identifiers]	4756
10	((passive or conventional or traditional) adj (teaching or instruction* or learning)).tw.	7798
11	(teaching adj (aid or aids or demonstration*)).tw.	1614
12	((closed or open ended) adj question*).tw.	4573
13	(socratic adj (strateg* or method* or technique* or approach*)).tw.	243
14	((flipped or inverted) adj classroom*).tw.	649
15	((problem based or inquiry based or enquiry based or exploration based or project based or active) adj (education or learning or teaching or training)).tw.	14584
16	((problem based or inquiry based or enquiry based or exploration based or project based) adj (technique* or method* or approach* or strateg*)).tw.	600
17	(problem adj1 solving).tw.	49840
18	(critical adj (think* or reflect*)).tw.	20145
19	(concept adj (learning or mapping)).tw.	3251
20	((outcome based or competency based or results based or performance based or skills based) adj (education or learning or teaching or training)).tw.	10024
21	((explicit or detailed) adj (instruction* or teaching)).tw.	2189
22	((diagnostic or prescriptive or precision) adj teaching).tw.	1961
23	(common adj (fallacy or fallacies or error* or blunder* or mistake* or misconception* or preconception*)).tw.	1042
24	((collaborative or cooperative or co operative or team) adj learning).tw.	18805
25	((mixed age or cross age) adj teaching).tw.	800
26	((mixed age or cross age or ability or within class or heterogeneous or homogenous) adj group*).tw.	4517
27	((individual* or differentiated or personalized or programmed or anchored) adj instruction*).tw.	19235
28	((child controlled or student controlled or learner controlled) adj (teaching or education or learning or instruction*)).tw.	1624
29	((child centered or student centered or learner centered or self paced or selfpaced or tailored) adj (education or teaching or training or learning or instruction*)).tw.	4026
30	((individual or independent or mastery or open or classroom or contract) adj learning).tw.	8329
31	(journal* adj1 writing).tw.	4123
32	((online or on line or web based or computer based or mobile or interactive) adj (educa-	18154

	tion or teaching or training or learning or instruction*))).tw.	
33	(teacher adj1 training).tw.	13661
34	or/1-33 [ERIC index termer + text words (generelle og spesielle) fra linje 2-33]	798310
35	(systematic review or meta-analysis or metaanalysis).mp. [SR]	7199
36	34 and 35	3090
37	("2000" or "2001" or "2002" or "2003" or "2004" or "2005" or "2006" or "2007" or "2008" or "2009" or "2010" or "2011" or "2012" or "2013" or "2014" or "2015" or "2016" or "2017" or "2018" or "2019" or "2020").yr. [Riktig år-limit fra 2000 til i dag?]	716834
38	36 and 37	2488

Appendix 4. SUPPORT Summaries checklist for making judgements about how much confidence to place in a systematic review

Review:
Assessed by:
Date:

Section A: Methods used to identify, include and critically appraise studies

<p>A.1 Were the criteria used for deciding which studies to include in the review reported? Did the authors specify:</p> <ul style="list-style-type: none"> <input type="checkbox"/> Types of studies <input type="checkbox"/> Participants <input type="checkbox"/> Intervention(s) <input type="checkbox"/> Outcome(s) <p><i>Coding guide - check the answers above YES: All four should be yes</i></p>	<ul style="list-style-type: none"> <input type="checkbox"/> Yes <input type="checkbox"/> Can't tell/partially <input type="checkbox"/> No
<p><i>Comments (note important limitations or uncertainty)</i></p>	
<p>A.2 Was the search for evidence reasonably comprehensive? Were the following done:</p> <ul style="list-style-type: none"> <input type="checkbox"/> Language bias avoided (no restriction of inclusion based on language) <input type="checkbox"/> No restriction of inclusion based on publication status <input type="checkbox"/> Relevant databases searched (including Medline + Cochrane Library) <input type="checkbox"/> Reference lists in included articles checked <input type="checkbox"/> Authors/experts contacted <p><i>Coding guide - check the answers above: YES: All five should be yes PARTIALLY: Relevant databases and reference lists are both ticked off</i></p>	<ul style="list-style-type: none"> <input type="checkbox"/> Yes <input type="checkbox"/> Can't tell/partially <input type="checkbox"/> No
<p><i>Comments (note important limitations or uncertainty)</i></p>	

<p>A.3 Is the review reasonably up-to-date? <i>Were the searches done recently enough that more recent research is unlikely to be found or to change the results of the review?</i></p> <p><i>Coding guide – consider how many years since the last search (e.g. if more than 10 years the review is unlikely to be up-to-date) and whether there is ongoing research</i></p>	<input type="checkbox"/> Yes <input type="checkbox"/> Can't tell/not sure <input type="checkbox"/> No
<p><i>Comments (note important limitations or uncertainty)</i></p>	
<p>A.4 Was bias in the selection of articles avoided? Did the authors specify:</p> <input type="checkbox"/> Explicit selection criteria <input type="checkbox"/> Independent screening of full text by at least 2 reviewers <input type="checkbox"/> List of included studies provided <input type="checkbox"/> List of excluded studies provided <i>Coding guide - check the above</i> <i>YES: All four should be yes</i>	<input type="checkbox"/> Yes <input type="checkbox"/> Can't tell/partially <input type="checkbox"/> No
<p><i>Comments (note important limitations or uncertainty)</i></p>	
<p>A.5 Did the authors use appropriate criteria to assess the risk for bias in analysing the studies that are included?[†] (See Appendix for an example of criteria - Assessing Risk of Bias Criteria for EPOC Reviews)</p> <input type="checkbox"/> The criteria used for assessing the risk of bias were reported <input type="checkbox"/> A table or summary of the assessment of each included study for each criterion was reported <input type="checkbox"/> Sensible criteria were used that focus on the risk of bias (and not other qualities of the studies, such as precision or applicability) <i>Coding guide - check the above</i> <i>YES: All four should be yes</i>	<input type="checkbox"/> Yes <input type="checkbox"/> Can't tell/partially <input type="checkbox"/> No
<p><i>Comments (note important limitations or uncertainty)</i></p>	
<p>A.6 Overall – how would you rate the methods used to identify, include and critically appraise studies? <i>Summary assessment score A relates to the 5 questions above.</i> <i>If the “No” or “Partial” option is used for any of the questions above, the review is likely to have important limitations.</i> <i>Examples of fatal flaws might include not reporting explicit selection criteria, not providing a list of included studies or not assessing the risk of bias in included studies.</i></p>	<input type="checkbox"/> Fatal flaws (limitations that are important enough that the results of the review are not reliable and they should not be used in the policy brief) <input type="checkbox"/> Important limitations (limitations that are important enough that it would be worthwhile to search for another systematic review and to interpret the results of this review cautiously, if a better review cannot be found) <input type="checkbox"/> Reliable (only minor limitations)
<p><i>Comments (note any fatal flaws or important limitations).</i></p>	

Section B: Methods used to analyse the findings

<p>B.1 Were the characteristics and results of the included studies reliably reported? Was there: <input type="checkbox"/> Independent data extraction by at least 2 reviewers <input type="checkbox"/> A table or summary of the characteristics of the participants, interventions and outcomes for the included studies <input type="checkbox"/> A table or summary of the results of the included studies. <i>Coding guide - check the answers above</i> <i>YES: All three should be yes</i></p>	<input type="checkbox"/> Yes <input type="checkbox"/> Partially <input type="checkbox"/> No <input type="checkbox"/> Not applicable (e.g. no included studies)
<p><i>Comments (note important limitations or uncertainty)</i></p>	
<p>B.2 Were the methods used by the review authors to analyse the findings of the included studies reported?</p>	<input type="checkbox"/> Yes <input type="checkbox"/> Partially <input type="checkbox"/> No <input type="checkbox"/> Not applicable (e.g. no studies or no data)
<p><i>Comments (note important limitations or uncertainty)</i></p>	
<p>B.3 Did the review describe the extent of heterogeneity? <input type="checkbox"/> Did the review ensure that included studies were similar enough that it made sense to combine them, sensibly divide the included studies into homogeneous groups, or sensibly conclude that it did not make sense to combine or group the included studies? <input type="checkbox"/> Did the review discuss the extent to which there were important differences in the results of the included studies? <input type="checkbox"/> If a meta-analysis was done, was the I², chi square test for heterogeneity or other appropriate statistic reported?</p>	<input type="checkbox"/> Yes <input type="checkbox"/> Can't tell/partially <input type="checkbox"/> No <input type="checkbox"/> Not applicable (e.g. no studies or no data)
<p><i>Comments (note important limitations or uncertainty)</i></p>	

<p>B.4 Were the findings of the relevant studies combined (or not combined) appropriately relative to the primary question the review addresses and the available data?</p> <p>How was the data analysis done?</p> <ul style="list-style-type: none"> <input type="checkbox"/> Descriptive only <input type="checkbox"/> Vote counting based on direction of effect <input type="checkbox"/> Vote counting based on statistical significance <input type="checkbox"/> Description of range of effect sizes <input type="checkbox"/> Meta-analysis <input type="checkbox"/> Meta-regression <input type="checkbox"/> Other: specify <input type="checkbox"/> Not applicable (e.g. no studies or no data) <p>How were the studies weighted in the analysis?</p> <ul style="list-style-type: none"> <input type="checkbox"/> Equal weights (this is what is done when vote counting is used) <input type="checkbox"/> By quality or study design (this is rarely done) <input type="checkbox"/> Inverse variance (this is what is typically done in a meta-analysis) <input type="checkbox"/> Number of participants <input type="checkbox"/> Other, specify: <input type="checkbox"/> Not clear <input type="checkbox"/> Not applicable (e.g. no studies or no data) <p>Did the review address unit of analysis errors?</p> <ul style="list-style-type: none"> <input type="checkbox"/> Yes - took clustering into account in the analysis (e.g. used intra-cluster correlation coefficient) <input type="checkbox"/> No, but acknowledged problem of unit of analysis errors <input type="checkbox"/> No mention of issue <input type="checkbox"/> Not applicable - no clustered trials or studies included <p><i>Coding guide - check the answers above</i></p> <p><i>If narrative OR vote counting (where quantitative analyses would have been possible) OR inappropriate table, graph or meta-analyses OR unit of analyses errors not addressed (and should have been) the answer is likely NO.</i></p> <p><i>If appropriate table, graph or meta-analysis AND appropriate weights AND the extent of heterogeneity was taken into account, the answer is likely YES.</i></p> <p><i>If no studies/no data: NOT APPLICABLE</i></p> <p><i>If unsure: CAN'T TELL/PARTIALLY</i></p>	<ul style="list-style-type: none"> <input type="checkbox"/> Yes <input type="checkbox"/> Can't tell/partially <input type="checkbox"/> No <input type="checkbox"/> Not applicable (e.g. no studies or no data)
<p><i>Comments (note important limitations or uncertainty)</i></p>	
<p>B.5 Did the review examine the extent to which specific factors might explain differences in the results of the included studies?</p> <ul style="list-style-type: none"> <input type="checkbox"/> Were factors that the review authors considered as likely explanatory factors clearly described? <input type="checkbox"/> Was a sensible method used to explore the extent to which key factors explained heterogeneity? <ul style="list-style-type: none"> <input type="checkbox"/> Descriptive/textual <input type="checkbox"/> Graphical <input type="checkbox"/> Meta-regression <input type="checkbox"/> Other 	<ul style="list-style-type: none"> <input type="checkbox"/> Yes <input type="checkbox"/> Can't tell/partially <input type="checkbox"/> No <input type="checkbox"/> Not applicable (e.g. too few studies, no important differences in the results of the included studies, or the included studies were so dissimilar that it would not make sense to explore heterogeneity of the results)
<p><i>Comments (note important limitations or uncertainty)</i></p>	

<p>B.6 Overall - how would you rate the methods used to analyse the findings relative to the primary question addressed in the review?</p> <p><i>Summary assessment score B relates to the 5 questions in this section, regarding the analysis.</i></p> <p><i>If the "No" or "Partial" option is used for any of the 5 preceding questions, the review is likely to have important limitations.</i></p> <p><i>Examples of fatal flaws might include not reporting critical characteristics of the included studies or not reporting the results of the included studies.</i></p>	<p><input type="checkbox"/> Fatal flaws (limitations that are important enough that the results of the review are not reliable and they should not be used in the policy brief)</p> <p><input type="checkbox"/> Important limitations (limitations that are important enough that it would be worthwhile to search for another systematic review and to interpret the results of this review cautiously, if a better review cannot be found)</p> <p><input type="checkbox"/> Reliable (only minor limitations)</p>
<p><i>Use comments to specify if relevant, to flag uncertainty or need for discussion</i></p>	

Section C: Overall assessment of the reliability of the review

<p>C.1 Are there any other aspects of the review not mentioned before which lead you to question the results?</p>	<p><input type="checkbox"/> Additional methodological concerns</p> <p><input type="checkbox"/> Robustness</p> <p><input type="checkbox"/> Interpretation</p> <p><input type="checkbox"/> Conflicts of interest (of the review authors or for included studies)</p> <p><input type="checkbox"/> Other</p> <p><input type="checkbox"/> No other quality issues identified</p>
<p>C.2 Based on the above assessments of the methods how would you rate the reliability of the review?</p> <p><input type="checkbox"/> Fatal flaws (exclude); briefly (and politely) state the reasons for excluding the review by completing the following sentence: <i>This review was not included in this policy brief for the following reasons:</i></p> <p><i>Comments (briefly summarise any key messages or useful information that can be drawn from the review for policy makers or managers):</i></p> <p><input type="checkbox"/> Important limitations; briefly (and politely) state the most important limitations by editing the following sentence, if needed, and specifying what the important limitations are: <i>This review has important limitations.</i></p> <p><input type="checkbox"/> Reliable; briefly note any comments that should be noted regarding the reliability of this review by editing the following sentence, if needed: <i>This is a good quality systematic review with only minor limitations.</i></p>	

NOTES

† **Risk of bias** is the extent to which bias may be responsible for the findings of a study.

Bias is a systematic error or deviation from the truth in results or inferences. In studies of the effects of health care, the main types of bias arise from systematic differences in the groups that are compared (selection bias), the care that is provided, or exposure to other factors apart from the intervention of interest (performance bias), withdrawals or exclusions of people entered into a study (attrition bias) or how outcomes are assessed (detection bias). Reviews of studies may also be particularly affected by reporting bias, where a biased subset of all the relevant data is available.

Assessments of the risk of bias are sometimes also referred to as assessments of the **validity** or **quality** of a study.

Validity is the extent to which a result (of a measurement or study) is likely to be true.

Quality is a vague notion of the strength or validity of a study, often indicating the extent of control over bias.

Suggested risk of bias criteria for EPOC reviews

Risk of bias for studies with a separate control group

Randomised controlled trials (RCTs)

Non-randomised controlled trials (NRCTs)

Controlled before-after (CBA) studies

Nine standard criteria are suggested for all RCTs, NRCTs and CBA studies. Further information can be obtained from the Cochrane handbook section on risk of bias.

Was the allocation sequence adequately generated?

Score "Yes" if a random component in the sequence generation process is described (eg Referring to a random number table). Score "No" when a non-random method is used (eg performed by date of admission). NRCTs and CBA studies should be scored "No". Score "unclear" if not specified in the paper.

Was the allocation adequately concealed?

Score "Yes" if the unit of allocation was by institution, team or professional and allocation was performed on all units at the start of the study; or if the unit of allocation was by patient or episode of care and there was some form of centralised randomisation scheme, an on-site computer system or sealed opaque envelopes were used. CBA studies should be scored "No". Score "unclear" if not specified in the paper.

Were baseline outcome measurements similar?^{1,2}

Score "Yes" if performance or patient outcomes were measured prior to the intervention, and no important differences were present across study groups. In RCTs, score "Yes" if imbalanced but appropriate adjusted analysis was performed (e.g. Analysis of covariance). Score "No" if important differences were present and not adjusted for in analysis. If RCTs have no baseline measure of outcome, score "Unclear".

Were baseline characteristics similar?

Score "Yes" if baseline characteristics of the study and control providers are reported and similar. Score "Unclear" if it is not clear in the paper (e.g. characteristics are mentioned in text but no data were presented). Score "No" if there is no report of characteristics in text or tables or if there are differences between control and intervention providers. Note that in some cases imbalance in patient characteristics may be due to recruitment bias whereby the provider was responsible for recruiting patients into the trial.

Were incomplete outcome data adequately addressed?¹

Score "Yes" if missing outcome measures were unlikely to bias the results (e.g. the proportion of missing data was similar in the intervention and control groups or the proportion of missing data was less than the effect size i.e. unlikely to overturn the study result). Score "No" if missing outcome data was likely to bias the results. Score "Unclear" if not specified in the paper (Do not assume 100% follow up unless stated explicitly).

Was knowledge of the allocated interventions adequately prevented during the study?¹

Score "Yes" if the authors state explicitly that the primary outcome variables were assessed blindly, or the outcomes are objective, e.g. length of hospital stay. Primary outcomes are those variables that correspond to the primary hypothesis or question as defined by the authors. Score "No" if the outcomes were not assessed blindly. Score "unclear" if not specified in the paper.

Was the study adequately protected against contamination?

Score "Yes" if allocation was by community, institution or practice and it is unlikely that the control group received the intervention. Score "No" if it is likely that the control group received the intervention (e.g. if patients rather than professionals were randomised). Score "unclear" if professionals were allocated within a clinic or practice and it is possible that communication between intervention and control professionals could have occurred (e.g. physicians within practices were allocated to intervention or control)

¹ If some primary outcomes were imbalanced at baseline, assessed blindly or affected by missing data and others were not, each primary outcome can be scored separately.

² If "UNCLEAR" or "No", but there is sufficient data in the paper to do an adjusted analysis (e.g. Baseline adjustment analysis or Intention to treat analysis) the criteria should be re scored as "Yes".

Was the study free from selective outcome reporting?

Score "Yes" if there is no evidence that outcomes were selectively reported (e.g. all relevant outcomes in the methods section are reported in the results section). Score "No" if some important outcomes are subsequently omitted from the results. Score "unclear" if not specified in the paper.

Was the study free from other risks of bias?

Score "Yes" if there is no evidence of other risk of biases

Risk of bias for interrupted time series (ITS) studies

Seven standard criteria are used for all ITS studies. Further information can be obtained from the Cochrane handbook section on Risk of Bias and from the draft methods paper on risk of bias under the EPOC specific resources section of the EPOC website.

Note: If the ITS study has ignored secular (trend) changes and performed a simple t-test of the pre versus post intervention periods without further justification, the study should not be included in the review unless reanalysis is possible.

Was the intervention independent of other changes?

Score "Yes" if there are compelling arguments that the intervention occurred independently of other changes over time and the outcome was not influenced by other confounding variables/historic events during study period. If Events/variables identified, note what they are. Score "NO" if reported that intervention was not independent of other changes in time.

Was the shape of the intervention effect pre-specified?

Score "Yes" if point of analysis is the point of intervention OR a rational explanation for the shape of intervention effect was given by the author(s). Where appropriate, this should include an explanation if the point of analysis is NOT the point of intervention; Score "No" if it is clear that the condition above is not met

Was the intervention unlikely to affect data collection?

Score "Yes" if reported that intervention itself was unlikely to affect data collection (for example, sources and methods of data collection were the same before and after the intervention); Score "No" if the intervention itself was likely to affect data collection (for example, any change in source or method of data collection reported).

Was knowledge of the allocated interventions adequately prevented during the study?³

Score "Yes" if the authors state explicitly that the primary outcome variables were assessed blindly, or the outcomes are objective, e.g. length of hospital stay. Primary outcomes are those variables that correspond to the primary hypothesis or question as defined by the authors. Score "No" if the outcomes were not assessed blindly. Score "unclear" if not specified in the paper.

Were incomplete outcome data adequately addressed?³

Score "Yes" if missing outcome measures were unlikely to bias the results (e.g. the proportion of missing data was similar in the pre- and post-intervention periods or the proportion of missing data was less than the effect size i.e. unlikely to overturn the study result). Score "No" if missing outcome data was likely to bias the results. Score "Unclear" if not specified in the paper (Do not assume 100% follow up unless stated explicitly).

Was the study free from selective outcome reporting?

Score "Yes" if there is no evidence that outcomes were selectively reported (e.g. all relevant outcomes in the methods section are reported in the results section). Score "No" if some important outcomes are subsequently omitted from the results. Score "unclear" if not specified in the paper.

Was the study free from other risks of bias?

Score "Yes" if there is no evidence of other risk of biases.

e.g. should consider if seasonality is an issue (i.e. if January to June comprises the pre-intervention period and July to December the post, could the "seasons" have caused a spurious effect).

³ If some primary outcomes were assessed blindly or affected by missing data and others were not, each primary outcome can be scored separately.

Appendix 5. Guidelines for preparing SUPPORT Summaries

Updated 18 March 2013

Examples of completed Summaries can be found on the SUPPORT website: www.supportsummaries.org

Remember:

- The audience is policymakers and their support staff, not researchers. SUPPORT Summaries are stand alone documents. Language should be plain language.
 - Avoid unnecessary jargon.
 - Explain useful jargon.
 - Use terms consistently throughout the summary; e.g. clinics rather than health posts / health centres / health facilities; lay health workers rather than community health workers.
 - Use footnotes to explain terms, if appropriate.
 - If appropriate, suggest terms that should be added to a glossary on the SUPPORT Summaries website.
 - Explain what is meant by terms that may be used in different ways; e.g. quality of care
 - Do not use abbreviations in the text. If abbreviations need to be used in a table, spell out the abbreviation in parentheses after (e.g. in the title or a heading) or spell them out in a footnote.
 - Use measures of effect that can be understood easily. If this is not possible spell out and explain terms (e.g. odds ratio or standardised mean differences) and provide an interpretation
 - Use standard terminology for study designs (See [Suggested terminology for study designs](#))
 - Use plain language terms for describing quality of evidence consistently throughout the summary and avoid non-standard terms; e.g. 'solid evidence'
 - Remember to use simple, clear language. Short sentences are generally easier to read and understand than long sentences.
 - Ask someone who knows nothing about the topic to read the Summary and check that it is understandable and makes sense.
- Make sure that it is possible for someone who is not familiar with the topic of the Summary to understand
 - What the intervention is
 - What the problem is
 - What the most important outcomes are
 - What the findings are

Also, include something about the context of the research when this is critical to understanding the effects of the intervention(s) and how they might work in other contexts.
- Make sure the key messages
 - Are prefaced by a statement of what the problem is that the intervention(s) address, if this is not obvious.
 - Are limited to no more than 5 and fit on one page
 - Use standard plain language statements
 - Are informative
 - Are supported by the findings of the review or the Relevance section
 - Appropriately interpret uncertainty and make clear the reason for uncertainty; e.g. "No studies were found that met the inclusion criteria of this review. We are therefore uncertain of the effects of ..."
 - Do not include statements regarding the quality of evidence
- Don't use 'statistically significant' or 'not statistically significant'.
- Don't use tables other than Summary of Findings tables.
- Avoid repetition and stating the obvious.
- Make sure that all of the Sections of the Summary are completed before you submit it.

Title and key messages



[Month and year] - SUPPORT Summary of a systematic review

[Title]

[Short background Text]

Key messages

- [Text]
- [Text]
- [Text]
- [Text]
- [Text]

Month and year when the summary was last revised; e.g. August 2011

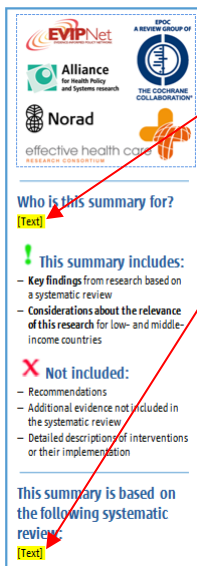
The title should be a question, such as: “Do lay health workers in primary and community health care improve maternal and child health?” or “What are the impacts of policies regarding direct patient payments for drugs?”

The background text here should not be more than 1 or 2 sentences and should only provide explanation that is essential to the understanding of the key messages.

There should not be more than 4 or 5 key messages summarising the most important messages from the summary of findings + a key message regarding the relevance of the review.

The key messages should not extend beyond page 1.

Key messages from the summary of findings should be phrased consistently with the messages in the summary of findings (using the plain language descriptions of findings on page 3). The quality of the evidence should not be included in the key messages on the first page.



Typically this should begin with “People making decisions concerning” followed by the topic of the review; e.g. “the use of conditional cash transfers to improve the uptake of health interventions.”

The citation should be in Vancouver style (e.g. Lagarde M, Haines A, Palmer N. Conditional cash transfers for improving uptake of health interventions in low and middle-income countries: a systematic review. *JAMA* 2007; 298:1900-10.) or recommended format for Cochrane reviews and other reports. **Check that it is correct! Include web address if the review is open access.**

Background

Background

[Text]

How this summary was prepared

After searching widely for systematic reviews that can help inform decisions about health systems, we have selected ones that provide information that is relevant to low- and middle-income countries. The methods used to assess the quality of the review and to make judgements about its relevance are described here: www.supportsummaries.org/methods

Knowing what's not known is important

A good quality review might not find any studies from low- and middle-income countries or might not find any well-designed studies. Although that is disappointing, it is important to know what is not known as well as what is known.

The background + the table below should be kept to one page. It should address **key important background information only** that is important to understand the objectives of the review, including explanation of which-ever of the following is not obvious or may be confusing, if not explained:

- The people, settings or problem
- The intervention(s) or policies
- The comparison
- The outcomes or goals of the interventions or policies

The text here should not repeat the descriptions of each comparison below or information on the first page or in the table.

About the systematic review underlying this summary

Review objective: [Text]

	What the review authors searched for	What the review authors found
Interventions	[Text]	[Text]
Participants	[Text]	[Text]
Settings	[Text]	[Text]
Outcomes	[Text]	[Text]
Date of most recent search: [Month and year]		
Limitations: [Text]		

[Citation]

Include the types of study designs that were looked for also.

Include something about study designs here and include the number of studies for each type of intervention OR study design.

If relevant include the numbers of different types of settings here; e.g. USA - Medi-caid/Medicare (7), city level (1), HMO (1), Preferred Provider Organisation (1), commercial plans (1), health insurance (1); Sweden - Public health insurance (1); Canada - drug program (3), health insurance program (2); Australia - Pharmaceutical benefits scheme (2); Nepal, - Health posts (1).

Include the countries (and number of studies per country) here.

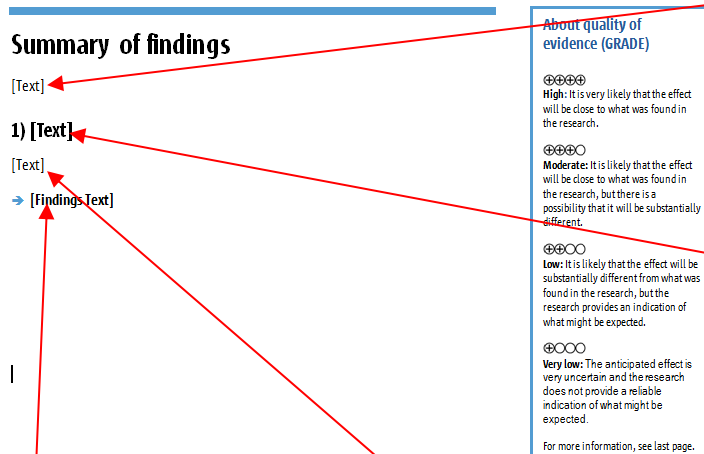
Include the number of studies that reported each primary out-come here.

This should either say: "This is well conducted systematic review with only minor limitations." OR be a brief statement of any important limitations. For example, "This is a well conducted systematic review with only minor limitations. However, it has not been updated since 1999." Or "This was an exhaustive review of English literature, but there were few evaluations of impact that allow robust conclusions to be drawn."

Do not say "with important limitations".

The citation at the bottom of this table should be the same as on the first page.

Summary of findings



The text here should be **one or two sentences** summarising the key information from the ‘about the review’ table; e.g. the total number of included studies and where they were done or the specific types of interventions for which studies were found (**NOT** the study designs).

If there is only one main comparison, this heading can be removed. Otherwise each numbered heading should specify a comparison or type of intervention.

As a rule, **the summary of findings should not exceed 3 pages** (up to 3 main comparisons with one comparison per page). Other comparisons in the review that are not included in the Summary can be listed at the end of the Summary of Findings.

The textual summary and Summary of Findings table for each comparison should be kept together on the same page.

The text here should be brief and should only provide information that is necessary to understand the specific comparison. It should not repeat what is in the background and should **NOT** include details about study designs.

The text here should correspond to and reflect what is in the table below. It is not necessary to use numbers here, if quantitative results are provided in the table and the quantitative results are not easily summarised or understood.

Use plain language expressions (See [Worksheets for preparing a summary of findings using GRADE.](#))

Include the quality of evidence together with these statements; e.g. Using lay health workers as an add-on to usual care probably increases immunisation coverage and breast feeding. The quality of this evidence was moderate.

Remember:

- **DO NOT SAY “no difference”!**
- **DO NOT USE “statistically significant”, “statistically non-significant”** or similar terms referring to statistical significance. (See [Results should not be reported as statistically significant or statistically non-significant.](#))
- The outcomes should be stated in plain language
- When reporting data, make it clear whether you are talking about rates, totals, proportions, etc.
- When reporting proportions, make it clear what they refer to (e.g. X% of.....)
- When reporting findings based on scales, the meaning of these scales needs to be explained
- Make sure any footnotes you insert are completed

[Text]				
People	[Text]			
Settings	[Text]			
Intervention	[Text]			
Comparison	[Text]			
Outcomes	Absolute effect*		Relative effect (95% CI)	Certainty of the evidence (GRADE)
	Without [text]	With [text]		
	Difference: [?] [text] per [?] [text] (Margin of error: [?] to [?] [text])			
[Text]	[?] per [?]	[?] per [?]	RR [?] ([?] to [?])	⊕○○○ Very low
	Difference: [?] [text] per [?] [text] (Margin of error: [?] to [?] [text])			
[Text]	[?] per [?]	[?] per [?]	RR [?] ([?] to [?])	⊕⊕○○ Low
	Difference: [?] [text] per [?] [text] (Margin of error: [?] to [?] [text])			
[Text]	[?] per [?]	[?] per [?]	RR [?] ([?] to [?])	⊕⊕⊕○ Moderate
	Difference: [?] [text] per [?] [text] (Margin of error: [?] to [?] [text])			
[Text]	[?] per [?]	[?] per [?]	RR [?] ([?] to [?])	⊕⊕⊕⊕ High
	Difference: [?] [text] per [?] [text] (Margin of error: [?] to [?] [text])			
[Text]	No included studies		-	-
<small>Margin of error = Confidence interval (95% CI) RR: Risk ratio GRADE: GRADE Working Group grades of evidence (see above and last page) * The risk WITHOUT the intervention is based on [text]. The corresponding risk WITH the intervention (and the 95% confidence interval for the difference) is based on the overall relative effect (and its 95% confidence interval).</small>				

This should succinctly describe the intervention or comparison

Include one SoF table for each included comparison – up to 3.

Do not include an SoF table if it does not provide any useful information (e.g. if no relevant studies were found).

Note what the baseline risk is based on.

Use whichever table format is best suited to how results are reported in the review.

Remove or add rows as needed.

Edit the column headings, if needed.

Include footnotes for any abbreviations or explanations that are needed.

Avoid using abbreviations when possible.

Remember to include the most important outcomes (even if there are no data) and not to

include more than seven outcomes.

See [SUPPORT Summary SoF worksheets](#) for examples and use those worksheets for preparing each SoF table.

[Text]		
People	[Text]	
Settings	[Text]	
Intervention	[Text]	
Comparison	[Text]	
Outcomes	Impact	Certainty of the evidence (GRADE)
[Text]	[Text]	⊕○○○ Very low
[Text]	[Text]	⊕⊕○○ Low
[Text]	[Text]	⊕⊕⊕○ Moderate
[Text]	[Text]	⊕⊕⊕⊕ High

Use this format if there is not a meta-analysis or if the results are reported in such a way that they cannot be summarised quantitatively in a consistent way for each outcome.

It may also be best to use this format if standardised mean differences or continuous outcomes that are not intuitively understandable are reported, stating the results in text in a way that is understandable; e.g. “Patients were on average more satisfied with care provided by a nurse practitioner than by a doctor (SMD +0.27, 95% CI +0.07 to +0.47)” So far as possible, avoid reporting SMD’s or other outcome measures that are difficult to understand and interpret.

So far as possible this should include quantitative information about the size of the effect. If this is not possible, the suggested plain language phrases (see page 3) should be used. If available, information about the range of effect sizes or confidence intervals should be included.

AVOID VOTE COUNTING! It also generally is not helpful (and may be misleading) to simply report an inventory of studies by reporting the results of each study one at a time. Options for reporting effects for interventions or groups of interventions when it does not make sense (or is not possible) to report an average effect across studies include reporting: plain language summaries, ranges or interquartile ranges.

Relevance of the review

- Make sure that the logic of your interpretations is clear under
- Try to convey only one message per bullet point
- Keep it simple. Avoid interpretations that are complicated

Relevance of the review for low-income countries

→ Findings	↳ Interpretation*
APPLICABILITY	
→ [Text]	↳ [Text]
EQUITY	
→ [Text]	↳ [Text]
ECONOMIC CONSIDERATIONS	
→ [Text]	↳ [Text]
MONITORING & EVALUATION	
→ [Text]	↳ [Text]

*Judgements made by the authors of this summary, not necessarily those of the review authors, based on the findings of the review and consultation with researchers and policymakers in low-income countries. For additional details about how these judgements were made see: www.supportsummaries.org/methods

This should be findings from the review (e.g. all of the studies were from high-income countries), NOT judgements.

These should be relevant judgements made based on the findings and, for example, what the review authors, you or others know about how the intervention works and how differences between the settings where the studies were done and settings in LMIC might modify the effectiveness or risks of the intervention(s).

Applicability

This should address the applicability of the findings in and across low-income countries.

Differences between health systems may result in a policy or programme option that is used in one setting not being feasible or acceptable in another. Or these differences may result in an option not working in the same way in another setting, or even achieving different impacts in another setting. A key challenge that policymakers and those supporting them must face is therefore the need to understand whether research evidence about an option can be applied to their setting. Systematic reviews make this task easier by summarising the evidence from studies conducted in a variety of different settings. Many systematic reviews, however, do not provide adequate descriptions of the features of the actual settings in which the original studies were conducted. The following questions can be used to guide assessments the applicability of the findings of a systematic review to low-income countries:

1. Were the studies included in the systematic review conducted in low-income settings or were the findings consistent across settings or time periods?
2. Are there important differences in on-the-ground realities and constraints that might substantially alter the feasibility and acceptability of the option?
3. Are there important differences in health system arrangements that may mean an option could not work in the same way?
4. Are there important differences in the baseline conditions that might yield different absolute effects even if the relative effectiveness was the same?
5. What insights can be drawn about options, implementation, and monitoring and evaluation? Even if there are reasonable grounds for concluding that the impacts of an option might differ in or across low-income countries or when there is little or no evidence, insights can be drawn from a systematic review about possible options, as well as approaches to the implementation of options and to monitoring and evaluation.

See: [SUPPORT Tools for evidence-informed health Policymaking \(STP\) 9: Assessing the applicability of the findings of a systematic review](#)

Relevance of the review for low-income countries

→ Findings	▷ Interpretation*
APPLICABILITY	
→ [Text]	▷ [Text]
EQUITY	
→ [Text]	▷ [Text]
ECONOMIC CONSIDERATIONS	
→ [Text]	▷ [Text]
MONITORING & EVALUATION	
→ [Text]	▷ [Text]

*Judgements made by the authors of this summary, not necessarily those of the review authors, based on the findings of the review and consultation with researchers and policymakers in low-income countries. For additional details about how these judgements were made see: www.supportsummaries.org/methods

Equity

This should address potential differences in effects for disadvantaged populations within countries: What impact is the policy or action likely to have on disadvantaged populations and equity in low-income countries?

Inequities can be defined as "differences in health which are not only unnecessary and avoidable but, in addition, are considered unfair and unjust". These have been well documented in relation to social and economic factors. Policies or programmes that are effective can improve the overall health of a population. However, the impact of such policies and programmes on inequities may vary: they may have no impact on inequities, they may reduce inequities, or they may exacerbate them, regardless of their overall effects on population health.

The following questions can be considered when making judgements about the potential impact a policy or programme option is likely to have on disadvantaged groups, and on equity in low-income countries:

1. Which groups or settings are likely to be disadvantaged in relation to the option being considered?
Consideration should be given to the following groups or settings:
 - **Economic status:** low-income populations are more likely to be responsive to changes in the prices of goods and services. Because they have less disposable income, tobacco tax increases, for example, could make such populations more likely to quit. But they would also be made more vulnerable as a result of having to spend more money on tobacco if they did not quit smoking
 - **Employment or occupation:** employer-funded insurance schemes may result in differences in coverage, with less coverage being likely for those who are unemployed, self-employed or employed in small companies
 - **Education:** school-based programmes would be expected to differentially affect those who attend versus those who do not attend schools. Information campaigns that rely on printed materials to improve the utilisation of health services might have differential impacts on illiterate or less-educated populations
 - **Place of residence:** access to care is commonly more difficult in rural areas. Any strategy, therefore, that does not take into account the need to improve the delivery of effective clinical or public health interventions is likely to be less effective in rural areas
 - **Gender:** strategies for involving stakeholders in priority setting may affect women and men differently, resulting in priorities that may have different impacts on women and men
 - **Ethnicity:** ethnic groups (e.g. those groups who consider themselves, or are considered by others, to share common characteristics which differentiate them from other groups in society) may have beliefs and attitudes relating to the acceptability of a particular policy or programme. Delivery strategies that do not take these perspectives into account are likely to be less effective amongst ethnic groups where an otherwise effective policy or programme might not be readily accepted
2. Are there plausible reasons for anticipating differences in the relative effectiveness of the option for disadvantaged groups or settings?

3. Are there likely to be different baseline conditions across groups or settings such that that the absolute effectiveness of the option would be different, and the problem more or less important, for disadvantaged groups or settings?
4. Are there important considerations that should be made when implementing the option in order to ensure that inequities are reduced, if possible, and that they are not increased?

See: [SUPPORT Tools for evidence-informed health Policymaking \(STP\) 10: Taking equity into consideration when assessing the findings of a systematic review](#)

Relevance of the review for low-income countries

→ Findings	▷ Interpretation*
APPLICABILITY	
→ [Text]	▷ [Text]
EQUITY	
→ [Text]	▷ [Text]
ECONOMIC CONSIDERATIONS	
→ [Text]	▷ [Text]
MONITORING & EVALUATION	
→ [Text]	▷ [Text]

*Judgements made by the authors of this summary, not necessarily those of the review authors, based on the findings of the review and consultation with researchers and policymakers in low-income countries. For additional details about how these judgements were made see: www.supportsummaries.org/methods

Economic considerations checklist

What are the most important economic consequences that will need to be considered when rolling out or scaling up the policy or action?

1. What are the most important economic consequences?

Examples of potentially important economic consequences that should be considered include:

1. Changes in use of healthcare resources
<ul style="list-style-type: none"> • Intervention <ul style="list-style-type: none"> ○ Human resources/time ○ Consumable supplies ○ Land, buildings, equipment • Additional (or fewer) hospitalisations, outpatient visits or home visits • Additional (or less) use of laboratory tests or examinations • Paid transportation (e.g. emergency transportation)
2. Changes in use of non-healthcare resources
<ul style="list-style-type: none"> • Home adaptation • Special diets • Transportation to healthcare facilities • Social services (e.g. housing, home assistance, occupational training) • Crime (e.g. theft, fraud, violence, police investigation, court costs)
3. Changes in use of patient and informal caregiver time
<ul style="list-style-type: none"> • Visits • Hospital admissions • Time of family or other informal caregivers
4. Changes in productivity
<ul style="list-style-type: none"> • Changes in productivity and the intrinsic value of changes in health status should be captured in the value or importance attached to health outcomes and should not be included as resource consequences.

2. Are there important considerations regarding the distribution of the costs and benefits of the intervention? Who pays should not determine whether resource consequences are considered (i.e. a broad 'societal' perspective should be taken). However, who pays and who benefits may be an important consideration with respect to equity.

3. Is there information about the total resource implications of expanding coverage of the intervention and sustaining it and what are the implications for scale up?

4. Is there important uncertainty about medium to long term economic consequences? The length of follow-up in the available studies may be an important consideration, if there is important uncertainty about longer term economic consequences.

5. Is there important uncertainty about the applicability of reported economic consequences?

If possible, important economic consequences should be considered in natural units in the summary rather than as monetary values, which cannot easily be applied across different settings since resource use (which may differ across settings) cannot be separated from unit costs (which are likely to differ across settings).

- The quality of evidence for economic consequences should be considered using the same (GRADE) criteria as those used for other impacts, if possible.
- Be cautious about reporting the results of cost-effectiveness or cost-utility analyses, since they often will not be applicable across different settings because of differences in resource use, unit costs, and the assumptions that are made. If only monetary values are reported or the results are limited or potentially misleading – leave out the results of cost-effectiveness analyses. Only include the results of cost-effectiveness analyses if they provide a good sense of the magnitude of the costs in relation-ship to the effects of the intervention that is likely to be similar across a range of LMIC settings.
- Do not make judgements about the balance between the net benefits and costs (whether an intervention is worth what it costs), but include any results and interpretation that could help decision-makers to do so.

See [SUPPORT Tools for evidence-informed health Policymaking \(STP\) 12: Finding and using research evidence about resource use and costs](#)

Relevance of the review for low-income countries

→ Findings	▷ Interpretation*
APPLICABILITY	
→ [Text]	▷ [Text]
EQUITY	
→ [Text]	▷ [Text]
ECONOMIC CONSIDERATIONS	
→ [Text]	▷ [Text]
MONITORING & EVALUATION	
→ [Text]	▷ [Text]

*Judgements made by the authors of this summary, not necessarily those of the review authors, based on the findings of the review and consultation with researchers and policymakers in low-income countries. For additional details about how these judgements were made see: www.supportsummaries.org/methods

Monitoring and evaluation

1. Is monitoring necessary?

The need for monitoring depends on the perceived need among relevant stakeholders to learn more about what is going on “on the ground”.

Whether it is worth the effort to set up a system for monitoring of a policy or programme may depend on several factors:

- Is there a monitoring system already in place that includes the needed indicators, or is a whole new set-up required?
- How much will it take to set up the required system? Is it as simple as adding a few items to data-collection procedures that already in place, or would additional large-scale household surveys be needed?
- Are the findings likely to be useful? What actions can or will be taken if monitoring reveals that things are not going as planned?

2. If monitoring is necessary, what should be measured?

Factors that need to be considered when selecting indicator(s) to collect for monitoring purposes include: validity, reproducibility, acceptability, feasibility, reliability, sensitivity to change, and predictive validity.

- In practice there will often be a trade-off between picking the optimal or desired indicators and having to accept the indicators which can be measured using existing data.
- There are good reasons not to select more indicators than needed: trying to limit the burden of data-collection being put on the health system, avoiding collection of data that are not utilised, and rather concentrate on collecting fewer data of high quality.

3. Is an impact evaluation necessary?

If there is insufficient evidence to be confident about the impacts of implementing a policy or action, the following should be considered. Positive answers to these questions suggest the need for well-designed field trials or “planned delays” in rolling out or scaling up an intervention.

- Is the intervention potentially ineffective or harmful?
- Are there important uncertainties about potentially important benefits, harms or costs (due to either the quality or applicability of the evidence)?
- Would evaluating the impact of the planned policy or action represent good value for money?
- Are the necessary resources for undertaking an impact evaluation likely to be available? If not, could they be obtained and would it be possible to collaborate with other countries?

4. If an impact evaluation is necessary, what should be evaluated and how?

- If a randomised trial is warranted and practical, what should be compared and what are the primary outcomes?
- If a randomised trial is not warranted or practical, what would be the optimal design and primary outcomes?

Implications for evaluation should be specific and they should be justified; i.e. what specific uncertainty should be addressed, how, and why addressing that uncertainty is important for people making decisions about an intervention (or how to address a problem) and key stakeholders. Statements such as “Evaluation is needed” are unhelpful and should not be made.

The following reasons for uncertainty can help to guide the types of research that might be needed:

By outcome for each of the most important outcomes	Possible implications for research
Study design	Need for randomised trials, if appropriate
Risk of bias	Need for better designed and executed studies
Inconsistency	Unexplained inconsistency: need for evaluation in relevant sub-groups
Indirectness	Need for studies that directly address the question of interest
Imprecision	Need for more studies with more participants

See [SUPPORT Tools for Evidence-informed Policymaking in health 18: Planning monitoring and evaluation of policies](#)

Additional information

Related literature

[References in review, ask authors or other experts, or perform search]

This summary was prepared by

[Authors, Institute, Country]

Conflict of interest

[Text]. For details, see: www.supportsummaries.org/coi

Acknowledgements

This summary has been peer reviewed by: [Name, Country;]

This review should be cited as

[Text]

The summary should be cited as

[Summary authors]. [Summary title]. A SUPPORT Summary of a systematic review. [Month and year]. www.supportsummaries.org

Keywords

All Summaries:

evidence-informed health policy, evidence-based, systematic review, health systems research, health care, low and middle-income countries, developing countries, primary health care



[Add additional summary-specific keywords. Will be taken out of the text and used only in the Properties field]

Include words and synonyms that people might use when searching for information that is found in this summary; including words that are not in the text.

This should include any relevant descriptors of:

- the population, setting, problem or condition
- the intervention(s)
- the comparison, if relevant
- outcomes, if relevant

(example text) This summary was prepared with additional support from:

	[Text]
	The South African Cochrane Centre , the only centre of the international Cochrane Collaboration in Africa, aims to ensure that health care decision making in Africa is informed by high quality, timely and relevant research evidence. www.mrc.ac.za/cochrane/cochrane.htm

Delete these, if not relevant. Insert any other logos + text that are appropriate.

This should include information that is helpful to understand the problem, provides details about the interventions, or helps to put the results of the review in a broader context. It can also include other relevant systematic reviews.

Check references used in the background and discussion of the review, ask the authors of the review and other experts in the area and, if relevant, search for specific types of background information that is relevant and cannot be found in the review.

Use full name (first name(s) followed by last name(s) without titles or organisations.

Typically: "None declared."

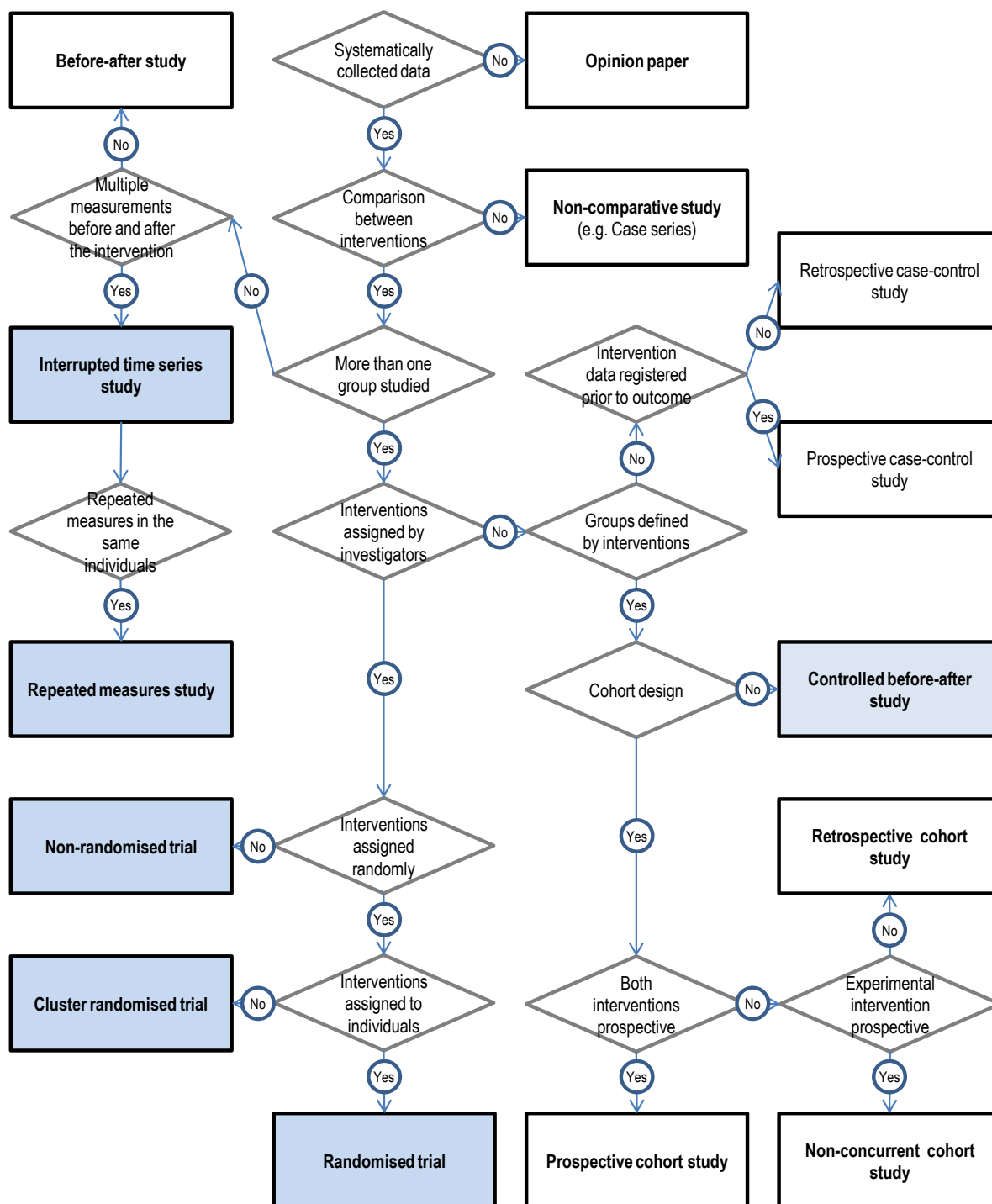
Suggested terminology for study designs

Suggested terms	Notes	Definition
Randomised controlled trial or preferably, randomised trial		An experimental study in which people are allocated to different interventions using methods that are random.
Non-randomised controlled trial OR preferably, non-randomised trial	Instead of controlled clinical trial . EPOC reviews do not include clinical trials (and RCTs are also CCTs). Also instead of ' quasi-randomised controlled trials ', which is used to mean different things by different authors.	An experimental study in which people are allocated to different interventions using methods that are not random.
Controlled before-after study	Instead of controlled before and after .	A study in which observations are made before and after the implementation of an intervention, both in a group that receives the intervention and in a control group that does not.
Interrupted-time-series study	Use study instead of design or analysis .	A study that uses observations at multiple time points before and after an intervention (the 'interruption'). The design attempts to detect whether the intervention has had an effect significantly greater than any underlying trend over time.
Repeated measures study		An ITS study where measurements are made in the same individuals at each time point.

For other study designs, use the terms in the algorithm below.

Study designs for evaluating the effects of healthcare interventions

(Shaded boxes are study designs that should be considered for inclusion in EPOC reviews.)



Appendix 6. Worksheets for preparing a summary of findings using GRADE

These worksheets can be used to:

1. Identify the most important outcomes for each comparison for which a SoF table would be helpful
2. Assess the quality of evidence for each of those outcomes using GRADE
3. Prepare a summary of findings (SoF) table for an EPOC review

Instructions

1. Identify each comparison in the review for which a SoF table would be helpful. Prepare more than one SoF table if the review contains more than one comparison for which a summary of findings would be helpful.
2. Select the most important outcomes for each comparison
Suggestions
 - a) Generate a list of relevant outcomes (see Worksheet 1)
 - List outcomes that you identified as primary outcomes
 - Add other outcomes for which data are reported
 - Add any other outcomes that were not reported in the review, but that might be important to someone making a decision – from the perspective of those who will be affected by the decision. **Be sure to consider potential benefits, adverse effects, and resource use (costs)**
 - Agree (with your co-authors) on which outcomes are important enough to be included in the SoF table (Worksheet 1)
 - b) Having chosen the outcomes that you think are most important and should be included in the SoF table, transfer them to a blank quality assessment table (see Worksheet 2).
 - Include outcomes that are critical to a decision even if the review does not provide any evidence
3. Assess the quality of evidence for each outcome using the GRADE approach
Suggestions
 - Fill in Worksheet 2 to determine the quality of the evidence for the outcome
 - Consult the criteria for assessing the quality of evidence (see below)
4. Summarise the findings for the outcome (quantitatively if possible), in a way that will be understandable to decision-makers and other stakeholders.
5. Complete the SoF table (Worksheet 3) filling in the Quality of the Evidence column for each of the important outcomes.
6. Prepare bullet points that summarise the information in the summary of findings table in plain language. Be consistent in how you translate the findings into qualitative statements (Worksheet 4) and your use of language when you report the results in the abstract, results, discussion and conclusions of the review.

Worksheet 1: Assessing the relative importance of outcomes and deciding which ones to include in the Summary of Findings table

Review:

Assessed by:

Date:

Rate the relative importance for each outcome on a 9 point scale ranging from 1 (not important) to 9 (critical).

1-3: Not important and not included in the SoF table

4-6: Important but not critical for making a decision (inclusion in the SoF table may depend on how many other important outcomes there are)

7-9: Critical for making a decision and should definitely be included in the SoF table

Include potential undesirable effects (harms) and resource use (costs), as well as desirable effects (benefits)

Outcome	Initials of people assessing the relative importance of the outcomes				Consensus
	Relative importance (1-9)				
a)					
b)					
c)					
d)					
e)					
f)					
g)					
h)					
i)					
j)					
k)					
l)					
m)					
n)					
o)					
p)					
q)					

Worksheet 2: Assessing the quality of evidence across studies for an outcome

(See the notes on quality assessment following the table below)

Comparison _____

Quality assessment of evidence for each outcome

No of studies	Design	Risk of bias	Inconsistency	Indirectness ⁴	Imprecision	Other ⁵	Quality (overall score) ⁶
Outcome:							
Outcome:							
Outcome:							
Outcome:							
Example: The use of lay health workers compared to usual health care services							
Outcome: Immunisation uptake in children							
4	Randomised trials (4)	Serious risk of bias (-0.5)	Important inconsistency (-0.5)	No serious indirectness	No serious imprecision	None	Moderate (3)

⁴ Indirectness includes consideration of

- Indirect (between study) comparisons
- Indirect (surrogate) outcomes
- Applicability (study populations, interventions or comparisons that are different than those of interest)

⁵ Other considerations for downgrading include publication bias. Other considerations for upgrading include a strong association with no plausible confounders, a dose response relationship, and if all plausible confounders or biases would decrease the size of the effect (if there is evidence of an effect), or increase it if there is evidence of no harmful effect (safety)

⁶ 4 ⊕⊕⊕⊕ **High** = We are confident that the true effect lies close to that of the estimate of the effect

3 ⊕⊕⊕⊖ **Moderate** = The true effect is likely to be close to the estimate of the effect, but there is a possibility that it is substantially different

2 ⊕⊕⊖⊖ **Low** = The true effect may be substantially different from the estimate of the effect

1 ⊕⊖⊖⊖ **Very low** = Any estimate of effect is very uncertain

Notes on quality assessment (scores generated in worksheet 2)

Quality of evidence assessment criteria

Quality of evidence	Study design	Lower if*	Higher if*
High (4)	Randomised trial	Study limitations (risk of bias) -1 Serious -2 Very serious	Strong association +1 Strong, no plausible confounders +2 Very strong, no major threats to validity
Moderate (3)		Inconsistency -1 Serious -2 Very serious	Dose response +1 Evidence of a gradient
Low (2)	Observational study	Indirectness -1 Serious -2 Very serious	All plausible confounders +1 All plausible confounders or bias would decrease the size of the effect if there is evidence of an effect, or increase it if there is evidence of no harmful effect (safety)
Very low (1)		Imprecision -1 Serious -2 Very serious Publication bias -1 Likely -2 Very likely	

* 1 = Move up or down one grade (for example from high to intermediate)

2 = Move up or down two grades (for example from high to low)

0.5 = Borderline

Generating scores for the quality of evidence across studies for an outcome involves making judgements about how much the factors in columns 3 and 4 decrease or increase the strength of the evidence. Details about the factors affecting the quality of the evidence can be found in the resources listed at the end of these worksheets.

You should include explanations for the judgements you made e.g. the evidence was downgraded from a high to moderate rating because of a risk of bias that borders on being serious (due perhaps to an incomplete follow-up or the absence of blinding in some of the trials) and an inconsistency of results across studies that borders on being important (ranging from inconclusive to a 36% relative increase).

Further guidance on generating quality of evidence scores, and a step by step guide to creating summary of findings tables can be found in GRADEpro, which can be downloaded from <http://ims.cochrane.org/revman/other-resources/gradepreresources>.

Worksheet 3: Summary of Findings (SoF) table

(Use the top rows for dichotomous outcomes when there is a meta-analysis. Use the bottom row for other outcomes.)

[Text]				
People	[Text]			
Settings	[Text]			
Intervention	[Text]			
Comparison	[Text]			
Outcomes	Absolute effect*		Relative effect (95% CI)	Certainty of the evidence (GRADE)
	Without lay health workers	With lay health workers		
[Text]	[?] per [?] Difference: [?] [Text] per [?] (95% CI: [?] to [?] [Text])	[?] per [?]	[?] [Text] (95% CI: [?] to [?] [Text])	⊕○○○ Very low
[Text]	[?] per 100 Difference: [?] [Text] per [?] (95% CI: [?] to [?] [Text])	[?] per 100	[?] [Text] (95% CI: [?] to [?] [Text])	⊕⊕○○ Low
[Text]	[?] per 100 Difference: [?] [Text] per [?] (95% CI: [?] to [?] [Text])	[?] per 100	[?] [Text] (95% CI: [?] to [?] [Text])	⊕⊕⊕○ Moderate
[Text]	[?] per 100 Difference: [?] [Text] per [?] (95% CI: [?] to [?] [Text])	[?] per 100	[?] [Text] (95% CI: [?] to [?] [Text])	⊕⊕⊕⊕ High
[Text]	[Text]		-	⊕⊕⊕○ Moderate

CI: Confidence interval RR: Risk ratio GRADE: GRADE Working Group grades of evidence (see above and last page)

* The risk WITHOUT the intervention is based on [Text]. The corresponding risk WITH the intervention (and the 95% confidence interval for the difference) is based on the overall relative effect (and its 95% confidence interval).

(Use this format if there is not a meta-analysis or if the results are reported in such a way that they cannot be summarised quantitatively in a consistent way for each outcome.)

[Text]		
People	[Text]	
Settings	[Text]	
Intervention	[Text]	
Comparison	[Text]	
Outcomes	Impact	Certainty of the evidence (GRADE)
[Text]	[Text]	⊕○○○ Very low
[Text]	[Text]	⊕⊕○○ Low
[Text]	[Text]	⊕⊕⊕○ Moderate
[Text]	[Text]	⊕⊕⊕⊕ High
GRADE: GRADE Working Group grades of evidence (see above and last page)		

(Use this format if the results are reported in such a way that they can be summarised quantitatively in a consistent way for each outcome.)

[Text]			
People	[Text]		
Settings	[Text]		
Intervention	[Text]		
Comparison	[Text]		
Outcomes	[Text] [Text]	Certainty of the evidence (GRADE)	Comments
[Text]	[Text]	⊕○○○ Very low	[Text]
[Text]	[Text]	⊕⊕○○ Low	[Text]
[Text]	[Text]	⊕⊕⊕○ Moderate	[Text]
[Text]	[Text]	⊕⊕⊕⊕ High	[Text]
GRADE: GRADE Working Group grades of evidence (see above and last page)			

(Use this format if the results are reported in such a way that they can be summarised quantitatively in a consistent way for each outcome and comments are not needed.)








[Text]		
People	[Text]	
Settings	[Text]	
Intervention	[Text]	
Comparison	[Text]	
Outcomes	[Text] [Text]	Certainty of the evidence (GRADE)
[Text]	[Text]	⊕○○○ Very low
[Text]	[Text]	⊕⊕○○ Low
[Text]	[Text]	⊕⊕⊕○ Moderate
[Text]	[Text]	⊕⊕⊕⊕ High
GRADE: GRADE Working Group grades of evidence (see above and last page)		

Summary of Findings – Examples

1. Summary of Findings – Substitution of nurses for physicians in primary care

Substitution of nurses for physicians in primary care		
People	All presenting patients in primary care	
Settings	Primarily Canada, the United States of America (USA) and the United Kingdom (UK)	
Intervention	Substitution of nurses for physicians (nurse-led primary care)	
Comparison	Routine care provided by physicians (physician-led primary care)	
Outcomes	Impact	Certainty of the evidence (GRADE)
Patient outcomes	The care provided by nurses and physicians may lead to similar health outcomes for patients.	⊕⊕○○ Low
Quality of care	The extent to which care provided by nurses was more or less appropriate than the care provided by physicians was not reported.	-
Patient satisfaction	On average patients are probably more satisfied with care provided by nurses, but some prefer care provided by nurses, and some prefer care provided by doctors.	⊕⊕⊕○ Moderate
Direct costs	The lower salary costs of nurses may be offset by their increased use of resources or lower productivity so that there may be little if any difference in the cost of care provided by nurses compared to the cost of care provided by physicians. Because the difference in salary between nurses and doctors may vary from place to place and over time, the net saving, if any, is likely to depend on the context.	⊕○○○ Very low
GRADE: GRADE Working Group grades of evidence (see above and last page)		



2. Summary of Findings – Lay health workers as an add on to usual care

Lay health workers as an add on to usual care				
People	Mothers or children under five			
Settings	Mixed (high-income countries for immunisations, mixed for breast feeding, low-income countries for morbidity and mortality in children)			
Intervention	Lay health workers (LHWs) (members of the community who are not health professionals and have received some training to promote health or to provide some health care services)			
Comparison	Usual care (varied across studies)			
Outcomes	Absolute effect*		Relative effect (95% CI)	Certainty of the evidence (GRADE)
	Without lay health workers	With lay health workers		
Mortality in children under five	5 per 100 Difference: 1 less per 100 (95% CI: (2 to 0 fewer))	4 per 100	25% relative reduction in child deaths (a 45% reduction to a 3% increase)	 Low
Neonatal mortality	4 per 100 Difference: 1 less per 100 (95% CI: (2 to 0 fewer))	3 per 100	24% relative reduction in infant deaths (a 43% reduction to a 2% increase)	 Low
Morbidity in children under five (e.g. fever, diarrhoea)	50 per 100 Difference: 7 fewer per 100 (95% CI: (13 to 1 fewer))	43 per 100	14% relative reduction in child morbidity (a 25 to a 1% reduction)	 Low
Care seeking for children under five	20 per 100 Difference: 7 more per 100 (95% CI: (3 fewer to 21 more))	27 per 100	33% relative increase in care seeking for children (a 14% reduction to a 105% increase)	 Low
Completed infant immunisations	45 per 100 Difference: 11 more per 100 (95% CI: (5 to 17 more))	55 per 100	22% relative increase in infant immunisations (a 10 to a 37% increase)	 Moderate
Initiation of breastfeeding	54 per 100 Difference: 18 more per 100 (95% CI: (7 to 33 more))	73 per 100	36% relative increase in initiated breast feeding (14 to 61%)	 Moderate
Exclusive breastfeeding	7 per 100 Difference: 16 more per 100 (95% CI: (5 to 24 more))	20 per 100	178% relative increase in exclusive breast feeding (74 to 344%)	 Moderate




CI: Confidence interval RR: Risk ratio GRADE: GRADE Working Group grades of evidence (see above and last page)

* The risk WITHOUT the intervention is based on the median control group risk across studies. The corresponding risk WITH the intervention (and the 95% confidence interval for the difference) is based on the overall relative effect (and its 95% confidence interval).

3. Summary of Findings – Educational meetings for health professionals

Educational meetings for health professionals			
People	Health care professionals		
Settings	Primary and secondary care		
Intervention	Educational meetings with or without other interventions*		
Comparison	No intervention		
Outcomes	Adjusted absolute improvement (risk difference)† Median (Interquartile range)	Certainty of the evidence (GRADE)	Comments
Compliance with desired practice	Median 6% (1.8 to 15.9)	 Moderate	The effect appears to be larger with higher attendance at the educational meetings and with mixed interactive and didactic educational meetings. Educational meetings did not appear to be effective for complex behaviours and they appeared to be less effective for less serious outcomes.
Patient outcomes	Median 3.0% (0.1% to 4.0%)	 Moderate	[Text]
GRADE: GRADE Working Group grades of evidence (see above and last page)			
* The effect of educational meetings alone on professional practice was the same as for multifaceted interventions that included educational meetings.			
†The post-intervention risk differences are adjusted for pre-intervention differences between the comparison groups.			

4. Summary of Findings – Introducing user fees

Introducing user fees			
People	Anyone using any type of health service in low- and middle-income countries		
Settings	Burkina Faso, Kenya, Lesotho, Papua New Guinea		
Intervention	Introducing or increasing user fees		
Comparison	No user fees		
Outcomes	Relative change in utilisation*	Certainty of the evidence (GRADE)	Comments
Healthcare utilisation – preventive care	15.4% less immediately 17% less after 12 months	 Very low	Antenatal care visits dropped in one study where fees were introduced. One additional study found a decrease in utilisation of deworming drugs following an introduction of fees, but did not report the results in a way that the relative change in utilisation could be calculated.
Healthcare utilisation – curative care	28% to 51% less immediately 9% less to 8% more after 12 months	 Very low	All but two studies showed a decrease in the number of outpatient visits in different types of facilities, although not all drops in attendance were statistically significant. Two controlled before-and-after studies where fees were introduced with quality improvements reported an increase in utilisation. However the authors did not report the results in a way that the relative change in utilisation could be calculated.
Equitable access – healthcare utilisation by quintile	N/A	 Very low	This study where quality improvements were introduced at the same time as user fees found an increase in utilisation for poor groups but not the very poorest (only quintiles 2 and 3). The authors did not report the results in a way that the relative change in utilisation could be calculated.
GRADE: GRADE Working Group grades of evidence (see above and last page)			
*For controlled before-after studies the relative change compared to the control group, and for interrupted time series studies the relative change compared to utilisation levels that would have been expected without the intervention			

Worksheet 4: Key messages in plain language

Prepare a small number of bullet points summarising the contents of the Summary of Findings table. Use consistent language, such as the following throughout the review. (Adapted from suggestions for Cochrane plain language summaries)

	Important difference	Small difference (May not be important)	Little or no difference
High quality evidence	Improves/decreases/prevents/ leads to [outcome]	Improves slightly/decreases slightly/leads to slightly fewer (more) [outcome]	Results in little or no difference in [outcome]
Moderate quality evidence	Probably improves/ decreases/ prevents/ leads to [outcome]	Probably improves slightly/decreases slightly/leads to slightly fewer (more) [outcome]	Probably leads to little or no difference in [outcome]
Low quality evidence	May improve/ decrease/prevent/lead to [outcome]	May slightly improve/slightly decrease/ lead to slightly fewer (more) [outcome]	May lead to little or no difference in [outcome]
Very low quality evidence	It is uncertain whether [intervention] improves, decreases, prevents, leads to [outcome] because the quality of the evidence is very low		
No data or no studies	[Outcome] was not measured or not reported, or no studies were found that evaluated the impact of [intervention] on [outcome]		

Plain language descriptions of the findings - Examples

Substitution of nurses for physicians in primary care (Example 1):

- Care provided by nurses and physicians may lead to similar health outcomes for patients
- It is uncertain whether there is any difference in the cost of care provided by nurses compared to the cost of care provided by physicians

Using lay health workers as an add-on to usual care (Example 2):

- Probably increases immunisation coverage and breast feeding
- May increase care seeking behaviour for children under five and reduce morbidity and mortality in children under five and neonates

Educational meetings for health professionals (Example 3):

- Probably improve compliance with desired practice and patient outcomes

Introducing user fees for health services in low- and middle-income countries (Example 4)

- It is uncertain whether introducing user fees reduces health service utilisation or increases inequities in low- and middle-income countries

Resources

Schünemann H, Brożek J, Oxman A, editors. GRADE handbook for grading quality of evidence and strength of recommendation. Version 3.2 [updated March 2009]. The GRADE Working Group, 2009. Available from www.cc-ims.net/gradepr.

Schünemann HJ, Oxman AD, Vist GE, Higgins JPT, Glasziou P, Guyatt GH. Presenting results and 'summary of findings' tables. Chapter 11. In: Higgins JPT, Green S (editors), *Cochrane Handbook for Systematic Reviews of Interventions*. Chichester (UK): Wiley-Blackwell, 2008.

Schünemann HJ, Oxman AD, Vist GE, Higgins JPT, Deeks JJ, Glasziou P, Guyatt GH. Interpreting results and drawing conclusions. Chapter 12. In: Higgins JPT, Green S (editors), *Cochrane Handbook for Systematic Reviews of Interventions*. Chichester (UK): Wiley-Blackwell, 2008.

Glenton C, Santesso N, Rosenbaum S, Nilsen ES, Rader T, Ciapponi A, H Dilkes. Presenting the results of Cochrane systematic reviews to a consumer audience: a qualitative study. *Med Decis Making*, 2010; 30:566-77.

Guyatt GH, Oxman AD, Kunz R, Vist GE, Falck-Ytter Y, Schunemann HJ, and the GRADE Working Group. What is 'quality of evidence' and why is it important to clinicians? *BMJ* 2008; 336:995-8.

Guyatt GH, Oxman AD, Akl EA, Kunz R, Vist G, Brozek J, et al. GRADE guidelines 1. Introduction - GRADE evidence profiles and summary of findings tables. *J Clin Epidemiol* 2011; 64:383-94.

Guyatt GH, Oxman AD, Kunz R, Atkins D, Brozek J, Vist G, et al. GRADE guidelines 2. Framing the question and deciding on important outcomes. *J Clin Epidemiol* 2011; 64: 395-400.

Balshem H, Helfand M, Schunemann H, Oxman AD, Kunz R, Brozek J, et al. GRADE guidelines 3. Rating the quality of evidence – introduction. *J Clin Epidemiol* 2011; 64:401-6.

Guyatt GH, Oxman AD, Vist G, Kunz R, Brozek J, Alonso-Coello P, et al. GRADE guidelines 4. Rating the quality of evidence – study limitations (risk of bias). *J Clin Epidemiol* 2011; 64:407-15.

Guyatt GH, Oxman AD, Montori V, Vist G, Kunz R, Brozek J, et al. GRADE guidelines - 5. Rating the quality of evidence - publication bias. *J Clin Epidemiol* 2011; 64:1277-82.

Guyatt GH, Oxman AD, Kunz R, Brozek J, Alonso-Coello P, Devereaux PJ, et al. GRADE guidelines 6. Rating the quality of evidence – imprecision. *J Clin Epidemiol* 2011; 64:1283-93.

Guyatt GH, Oxman AD, Kunz R, Woodcock J, Brozek J, Helfand M, et al. GRADE guidelines 7. Rating the quality of evidence – inconsistency. *J Clin Epidemiol* 2011; 64:1294-302.

Guyatt GH, Oxman AD, Kunz R, Woodcock J, Brozek J, Helfand M, et al. GRADE guidelines 8. Rating the quality of evidence – indirectness. *J Clin Epidemiol* 2011; 64:1303-10.

Guyatt GH, Oxman AD, Sultan S, Glasziou P, Alonso-Coello P, Atkins D, et al. GRADE guidelines 9. Rating up the quality of evidence. *J Clin Epidemiol* 2011; 64:1311-6.

Brunetti M, Shemilt I, Pregno S, Vale L, Oxman AD, Lord J, Sisk J, Francis R, Hill S, Guyatt GH, Jaeschke R, Helfand M, Harbour R, Davoli M, Amato L, Liberati A, Schünemann HJ. GRADE guidelines: 10. Considering resource use and rating the quality of economic evidence. *J Clin Epidemiol* 2013; 66:140-50.

Guyatt GH, Oxman AD, Sultan S, Glasziou P, Alonso-Coello P, Atkins D, Kunz R, Brozek J, Montori V, Jaeschke R, Rind D, Dahm P, Akl EA, Meerpohl J, Vist G, Berliner E, Norris S, Falck-ytter Y, Schünemann HJ. GRADE guidelines: 11. Making an overall rating of quality of evidence for a single outcome and for all outcomes. *J Clin Epidemiol* 2013; 66:151-7.

Guyatt GH, Oxman AD, Santesso N, Helfand M, Vist G, Kunz R, Brozek J, Norris S, Meerpohl J, Djulbecovic B, Alonso-Coello P, Post PN, Busse JW, Glasziou P, Christensen R, Schünemann HJ. GRADE guidelines: 12. Preparing summary of findings tables – binary outcomes. *J Clin Epidemiol*. 2013; 66:158-72.

Guyatt GH, Thorlund K, Oxman AD, Walter SD, Patrick D, Furukawa TA, Johnston BC, Karanicolas P, Akl EA, Vist G, Kunz R, Brozek J, Kupper LL, Martin SL, Meerpohl JJ, Alonso-Coello P, Christensen R, Schünemann HJ. GRADE guidelines: 13. Preparing summary of findings tables – continuous outcomes. *J Clin Epidemiol* 2013; 66:173-83.

Results should not be reported as statistically significant or statistically non-significant

Because “statistical significance” is so commonly misreported and misinterpreted, we recommend that the concept and related terms (e.g. not significant, not statistically significant, significant, statistically significant) should not be used in NsEPOC reviews. In general point estimates and confidence intervals, when possible, or p-values should be reported. Plain language should be used to describe effects based on the size of the effect and the quality of the evidence. (See [Worksheets for preparing summary of findings tables using GRADE.](#))

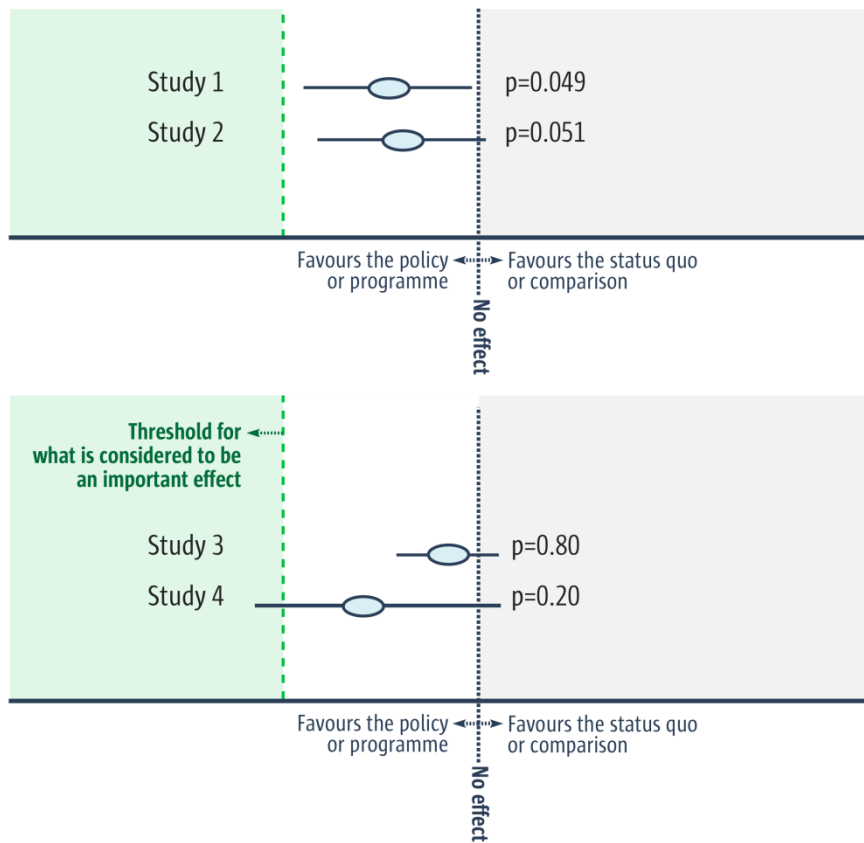
A common mistake made in instances when evidence is inconclusive is the confusion of a lack of evidence of an effect with ‘evidence of no effect’.⁷ It is wrong to claim that inconclusive evidence shows that an intervention has had ‘no effect’. ‘Statistical significance’ should *not* be confused with the size or importance of an effect.

When results are not ‘statistically significant’ it cannot be assumed that there was no impact. Typically a cut-off of 5% is used to indicate statistical significance. This means that the results are considered to be ‘statistically non-significant’ if the analysis shows that differences as large as (or larger than) the observed difference would be expected to occur by chance *more* than one out of twenty times ($p \geq 0.05$). There are, however, two problems with this assumption. Firstly, the cut-off point of 5% is arbitrary. Secondly, ‘statistically non-significant’ results (sometimes mislabelled as ‘negative’), might or might not be inconclusive. The figure below illustrates how the use of the terms ‘statistically non-significant’ or ‘negative’ can be misleading. Similarly, ‘statistically significant’ results might or might not be important.

Trends that are ‘positive’ (i.e. in favour of an option) but ‘statistically non-significant’ are often described as ‘promising’ and this can also be misleading. ‘Negative’ trends of the same magnitude, in contrast, are not typically described as ‘warning signs’.

⁷ Alderson P, Chalmers I: Survey of claims of no effect in abstracts of Cochrane reviews. *BMJ* 2003, 326:475.

Figure. Two problems with classifying results as 'statistically non-significant' or 'negative'



The blue dots in this figure indicate the estimated effect for each study and the horizontal lines indicate the 95% confidence intervals. A 95% confidence interval means that we can be 95 % confident that the true size of the effect is between the lower and upper confidence limit (the two ends of the horizontal lines). Conversely, there is a 5% chance that the true effect is outside this range.

The figure illustrates two problems that arise when results are classified as 'statistically non-significant' or 'negative':

1. **The classification is based on an arbitrary cut-off.** The results of Study 1, for example, are marginally different from the results of Study 2. But by using the conventional cut-off of $P < 0.05$, the results of Study 1 are considered 'statistically significant' and the results of Study 2 'statistically non-significant'
2. **'Statistically non-significant' results may or may not be inconclusive.** If the short green vertical line in the lower part of the figure indicates the smallest effect considered important, the results for Study 3 would be conclusive, since an important impact is highly unlikely. The results for Study 4 would be 'inconclusive' since it is not unlikely that there would be an important impact (the 95% confidence interval crosses the threshold for what is considered to be an important effect). Both results, however, might be regarded as 'statistically non-significant' or 'negative'

Appendix 7. SUPPORT Summary peer review form

SUPPORT summaries of systematic reviews are intended for people making decisions about health systems in low-income countries (LIC), particularly policymakers and managers (and their support staff). We would appreciate your advice regarding the following and any specific suggestions you have for improving the summary, particularly the **key messages** and the section of the summary that addresses the **relevance** of the review for LIC:

General

1. Do you consider the results of this review to be relevant to LIC?
2. Are you aware of another systematic review that addresses the same or a similar topic that is more up-to-date, of better quality, or more relevant to LIC?

Title, introduction and key messages

3. Are the key messages clear, relevant and consistent with the findings of the review?
4. Are there any changes that you would suggest to the title, brief introduction or key messages?

Background

5. Is the background clear, concise and appropriate for policymakers and managers in LIC?
6. Does the box describing the review ('About the systematic review underlying this summary') clearly and accurately reflect what the review authors searched for and found?

Summary of findings

7. Is the summary of the review's findings accurate, clear, relevant and appropriate for policymakers and managers in LIC?

Relevance

8. Are the interpretations that are made appropriate, relevant and useful for policymakers and managers in LIC?
9. Are there any additional comments or specific changes that you would suggest regarding applicability, equity, economic considerations, or monitoring and evaluation?
10. Are there comments regarding applicability, equity, economic considerations, or monitoring and evaluation that do not require a change to the summary, but that you would want to publish together with the revised summary? It will be possible for users of the summary to submit comments online, which will be published at the end of the review. Any relevant comments that you have would be most welcome and could be included when the summary is first published.

Additional information

11. Is there any other literature on the topic of this review that you think would be particularly useful to policymakers and managers in LIC, including related systematic reviews, information that is helpful to understand the problem, provides details about the interventions, or helps to put the results of the review in a broader context?
12. Is it OK to acknowledge you for peer reviewing this summary?

Appendix 8. Editor's checklist for review of SUPPORT Summaries

Summary:

Editor:

Date:

Checklist items	Suggestions	
<u>Is it possible for someone who is not familiar with the topic of the Summary to understand</u>		
• What the intervention is?	<input type="checkbox"/> Yes	<input type="checkbox"/> No
• What the problem is?	<input type="checkbox"/> Yes	<input type="checkbox"/> No
• What the most important outcomes are?	<input type="checkbox"/> Yes	<input type="checkbox"/> No
• What the findings are?	<input type="checkbox"/> Yes	<input type="checkbox"/> No
Are the <u>key messages</u>		
• Informative?	<input type="checkbox"/> Yes	<input type="checkbox"/> No
• Supported by the findings of the review (in the Summary) or the Relevance section?	<input type="checkbox"/> Yes	<input type="checkbox"/> No
• Expressed using appropriate plain language statements? *	<input type="checkbox"/> Yes	<input type="checkbox"/> No
• When there is no evidence, is this appropriately interpreted as uncertainty; e.g. "No studies were found that met the inclusion criteria of this review. We are therefore uncertain of the effects of ..."	<input type="checkbox"/> Yes	<input type="checkbox"/> No
In the <u>Summary of findings</u> section		
• Has use of 'statistical significance' or 'non-significance' been avoided? †	<input type="checkbox"/> Yes	<input type="checkbox"/> No
• Has lack of evidence been appropriately interpreted as uncertainty (rather than evidence of no effect)? †	<input type="checkbox"/> Yes	<input type="checkbox"/> No
• Study designs should not be included, unless necessary to understand the results.	<input type="checkbox"/> Yes	<input type="checkbox"/> No
In the <u>Relevance</u> section		
• Are the bullet points under Findings actually findings?	<input type="checkbox"/> Yes	<input type="checkbox"/> No
• Are the interpretations informative?	<input type="checkbox"/> Yes	<input type="checkbox"/> No
• If not, how might they be made more informative?	<input type="checkbox"/> Yes	<input type="checkbox"/> No
• Are suggestions for monitoring and evaluation specific/implementable?	<input type="checkbox"/> Yes	<input type="checkbox"/> No
Have all <u>unnecessary acronyms and jargon</u> been removed?	<input type="checkbox"/> Yes	<input type="checkbox"/> No
Should the Summary be <u>shortened</u> ?	<input type="checkbox"/> Yes	<input type="checkbox"/> No
<u>Other suggestions</u> for improving this Summary	<input type="checkbox"/> Yes	<input type="checkbox"/> No

* See SUPPORT Summary SoF worksheets 2013 05 09

† See [Results should not be reported as statistically significant or statistically non-significant](#))

Generic suggestions for improving the template, guidance or other Summaries:

Appendix 9. SUPPORT Summary template

[Month and year] – SUPPORT Summary of a systematic review

[Title]

[Short background Text]

Key messages

[Text]

→ [Text]

→ [Text]
[Text]



Who is this summary for?

[Text]

! This summary includes:

- **Key findings** from research based on a systematic review
- **Considerations about the relevance of this research** for low-income countries

X Not included:

- Recommendations
- Additional evidence not included in the systematic review
- Detailed descriptions of interventions or their implementation

This summary is based on the following systematic review:

[Text]

What is a systematic review?

A summary of studies addressing a clearly formulated question that uses systematic and explicit methods to identify, select, and critically appraise the relevant research, and to collect and analyse data from the included studies

SUPPORT was an international project to support the use of policy relevant reviews and trials to inform decisions about maternal and child health in low- and middle-income countries, funded by the European Commission (FP6) and the Canadian Institutes of Health Research.

Glossary of terms used in this report:

www.supportsummaries.org/glossary-of-terms

Background references on this topic:

See back page

Background

[Text]

How this summary was prepared

After searching widely for systematic reviews that can help inform decisions about health systems, we have selected ones that provide information that is relevant to low-income countries. The methods used to assess the reliability of the review and to make judgements about its relevance are described here:

www.supportsummaries.org/how-support-summaries-are-prepared/

Knowing what's not known is important

A reliable review might not find any studies from low-income countries or might not find any well-designed studies. Although that is disappointing, it is important to know what is not known as well as what is known.

A lack of evidence does not mean a lack of effects. It means the effects are uncertain. When there is a lack of evidence, consideration should be given to monitoring and evaluating the effects of the intervention, if it is used.

About the systematic review underlying this summary

Review objective: [Text]

Types of	What the review authors searched for	What the review authors found
Study designs & Interventions	[Text]	[Text]
Participants	[Text]	[Text]
Settings	[Text]	[Text]
Outcomes	[Text]	[Text]

Date of most recent search: [Month and year]

Limitations: [Text]

[Citation]

Summary of findings

[Text]

1) [Text]

[Text]

[Findings Text]

About the certainty of the evidence (GRADE) *



High: This research provides a very good indication of the likely effect. The likelihood that the effect will be substantially different† is low.



Moderate: This research provides a good indication of the likely effect. The likelihood that the effect will be substantially different† is moderate.



Low: This research provides some indication of the likely effect. However, the likelihood that it will be substantially different† is high.



Very low: This research does not provide a reliable indication of the likely effect. The likelihood that the effect will be substantially different† is very high.

* This is sometimes referred to as 'quality of evidence' or 'confidence in the estimate'.

† Substantially different = a large enough difference that it might affect a decision

See last page for more information.

(Use the top rows for dichotomous outcomes when there is a meta-analysis. Use the bottom row for other outcomes.)

[Text]				
People	[Text]			
Settings	[Text]			
Intervention	[Text]			
Comparison	[Text]			
Outcomes	Absolute effect*		Relative effect (95% CI)	Certainty of the evidence (GRADE)
	Without [text]	With [text]		
	Difference: [?] [text] per [?] [text] (Margin of error: [?] to [?] [text])			
[Text]	[?] per [?]	[?] per [?]	RR [?] ([?] to [?])	⊕⊕○○ Low
	Difference: [?] [text] per [?] [text] (Margin of error: [?] to [?] [text])			
[Text]	[?] per [?]	[?] per [?]	RR [?] ([?] to [?])	⊕⊕○○ Low
	Difference: [?] [text] per [?] [text] (Margin of error: [?] to [?] [text])			
[Text]	No included studies		-	-

Margin of error = Confidence interval (95% CI) RR: Risk ratio GRADE: GRADE Working Group grades of evidence (see above and last page)

* The risk WITHOUT the intervention is based on [Text]. The corresponding risk WITH the intervention (and the 95% confidence interval for the difference) is based on the overall relative effect (and its 95% confidence interval).

(Use this format if there is not a meta-analysis or if the results are reported in such a way that they cannot be summarised quantitatively in a consistent way for each outcome.)

[Text]		
People	[Text]	
Settings	[Text]	
Intervention	[Text]	
Comparison	[Text]	
Outcomes	Impact	Certainty of the evidence (GRADE)
[Text]	[Text]	⊕○○○ Very low
[Text]	[Text]	⊕⊕○○ Low
[Text]	[Text]	⊕⊕⊕○ Moderate
[Text]	[Text]	⊕⊕⊕⊕ High
GRADE: GRADE Working Group grades of evidence (see above and last page)		

(Use this format if the results are reported in such a way that they can be summarised quantitatively in a consistent way for each outcome.)

[Text]			
People	[Text]		
Settings	[Text]		
Intervention	[Text]		
Comparison	[Text]		
Outcomes	[Text] *	Certainty of the evidence (GRADE)	Comments
[Text]	[Text] [Text]	⊕○○○ Very low	[Text]
[Text]	[Text] [Text]	⊕⊕○○ Low	[Text]
[Text]	[Text] [Text]	⊕⊕⊕○ Moderate	[Text]
[Text]	[Text] [Text]	⊕⊕⊕⊕ High	[Text]
GRADE: GRADE Working Group grades of evidence (see above and last page)			
* [Text]			

(Use this format if the results are reported in such a way that they can be summarised quantitatively in a consistent way for each outcome and comments are not needed.)

[Text]		
People	[Text]	
Settings	[Text]	
Intervention	[Text]	
Comparison	[Text]	
Outcomes	[Text] * [Text]	Certainty of the evidence (GRADE)
[Text]	[Text] [Text]	⊕○○○ Very low
[Text]	[Text] [Text]	⊕⊕○○ Low
[Text]	[Text] [Text]	⊕⊕⊕○ Moderate
[Text]	[Text] [Text]	⊕⊕⊕⊕ High
GRADE: GRADE Working Group grades of evidence (see above and last page)		
* [Text]		

2) [Text]

[Text]

[Findings Text]

[Findings Text]

[Text]				
People	[Text]			
Settings	[Text]			
Intervention	[Text]			
Comparison	[Text]			
Outcomes	Absolute effect*		Relative effect (95% CI)	Certainty of the evidence (GRADE)
	Without [text]	With [text]		
[Text]	[?] per [?]	[?] per [?]	RR [?] ([?] to [?])	⊕⊕○○ Low
	Difference: [?] [text] per [?] [text] (Margin of error: [?] to [?] [text])			
[Text]	[?] per [?]	[?] per [?]	RR [?] ([?] to [?])	⊕⊕○○ Low
	Difference: [?] [text] per [?] [text] (Margin of error: [?] to [?] [text])			
[Text]	[?] per [?]	[?] per [?]	RR [?] ([?] to [?])	⊕⊕○○ Low
	Difference: [?] [text] per [?] [text] (Margin of error: [?] to [?] [text])			
[Text]	[?] per [?]	[?] per [?]	RR [?] ([?] to [?])	⊕⊕○○ Low
	Difference: [?] [text] per [?] [text] (Margin of error: [?] to [?] [text])			
[Text]	No included studies		-	-

Margin of error = Confidence interval (95% CI) RR: Risk ratio GRADE: GRADE Working Group grades of evidence (see above and last page)

* The risk WITHOUT the intervention is based on [Text]. The corresponding risk WITH the intervention (and the 95% confidence interval for the difference) is based on the overall relative effect (and its 95% confidence interval).

3) [Text]

[Text]

[Findings Text]

[Findings Text]

[Text]				
People	[Text]			
Settings	[Text]			
Intervention	[Text]			
Comparison	[Text]			
Outcomes	Absolute effect*		Relative effect (95% CI)	Certainty of the evidence (GRADE)
	Without [text]	With [text]		
[Text]	[?] per [?]	[?] per [?]	RR [?] ([?] to [?])	⊕⊕○○ Low
	Difference: [?] [text] per [?] [text] (Margin of error: [?] to [?] [text])			
[Text]	[?] per [?]	[?] per [?]	RR [?] ([?] to [?])	⊕⊕○○ Low
	Difference: [?] [text] per [?] [text] (Margin of error: [?] to [?] [text])			
[Text]	[?] per [?]	[?] per [?]	RR [?] ([?] to [?])	⊕⊕○○ Low
	Difference: [?] [text] per [?] [text] (Margin of error: [?] to [?] [text])			
[Text]	[?] per [?]	[?] per [?]	RR [?] ([?] to [?])	⊕⊕○○ Low
	Difference: [?] [text] per [?] [text] (Margin of error: [?] to [?] [text])			
[Text]	No included studies		-	-

Margin of error = Confidence interval (95% CI) RR: Risk ratio GRADE: GRADE Working Group grades of evidence (see above and last page)

* The risk WITHOUT the intervention is based on [Text]. The corresponding risk WITH the intervention (and the 95% confidence interval for the difference) is based on the overall relative effect (and its 95% confidence interval).

Relevance of the review for low-income countries

→ Findings	▷ Interpretation*
APPLICABILITY	
→ [Text]	▷ [Text]
EQUITY	
→ [Text]	▷ [Text]
ECONOMIC CONSIDERATIONS	
→ [Text]	▷ [Text]
MONITORING & EVALUATION	
→ [Text]	▷ [Text]

*Judgements made by the authors of this summary, not necessarily those of the review authors, based on the findings of the review and consultation with researchers and policymakers in low-income countries. For additional details about how these judgements were made see: www.supportsummaries.org/methods

Additional information

Related literature

[References in review, ask authors or other experts, or perform search]

This summary was prepared by

[Authors, Institute, Country]

Conflict of interest

[Text]. For details, see: www.supportsummaries.org/coi

Acknowledgements

This summary has been peer reviewed by: [Name, Country;]

This review should be cited as

[Text]

The summary should be cited as

[Summary authors]. [Summary title]. A SUPPORT Summary of a systematic review. [Month and year]. www.supportsummaries.org

Keywords

All Summaries:

evidence-informed health policy, evidence-based, systematic review, health systems research, health care, low and middle-income countries, developing countries, primary health care

[Add additional summary-specific keywords. Will be taken out of the text and used only in the Properties field]

(example text) This summary was prepared with additional support from:



[Text]



The South African Cochrane Centre, the only centre of the international Cochrane Collaboration in Africa, aims to ensure that health care decision making in Africa is informed by high quality, timely and relevant research evidence. www.mrc.ac.za/cochrane/cochrane.htm

About certainty of the evidence (GRADE)

The “certainty of the evidence” is an assessment of how good an indication the research provides of the likely effect; i.e. the likelihood that the effect will be substantially different from what the research found. By “substantially different” we mean a large enough difference that it might affect a decision. These judgements are made using the GRADE system, and are provided for each outcome. The judgements are based on the study design (randomised trials versus observational studies), factors that reduce the certainty (risk of bias, inconsistency, indirectness, imprecision, and publication bias) and factors that increase the certainty (a large effect, a dose response relationship, and plausible confounding). For each outcome, the certainty of the evidence is rated as high, moderate, low or very low using the definitions on page 3.

For more information about GRADE:
www.supportsummaries.org

SUPPORT collaborators:

The Cochrane Effective Practice and Organisation of Care Group (EPOC) is part of the **Cochrane Collaboration**. The Norwegian EPOC satellite supports the production of Cochrane reviews relevant to health systems in low- and middle-income countries.

www.epocoslo.cochrane.org

The Evidence-Informed Policy Network (EVIPNet) is an initiative to promote the use of health research in policymaking in low- and middle-income countries.

www.evipnet.org

The Alliance for Health Policy and Systems Research (HPSR) is an international collaboration that promotes the generation and use of health policy and systems research in low- and middle-income countries.

www.who.int/alliance-hpsr

Norad, the Norwegian Agency for Development Cooperation, supports the Norwegian EPOC satellite and the production of SUPPORT Summaries.

www.norad.no

The Effective Health Care Research Consortium is an international partnership that prepares Cochrane reviews relevant to low-income countries.