

Supplementary file 2 - Data extraction and coding tools

Coding tools containing full list of descriptive data to be extracted

A. Data extraction tool for impact evaluations

Category		Answer
Descriptive information	ID	Open answer – xxx
	Title	Open answer – this is the title that will be displayed on the EGM platform
	Full title	Open answer – this is the title of the paper
	Author Citation	Open answer
	Publication date	Open answer
	Map	Performance measurement and management in primary care systems in low and middle-income countries evidence gap map
	Regions (3ie)	<ul style="list-style-type: none"> – East Asia and Pacific – South Asia – Europe – CIS – Middle East and North Africa – Sub-Saharan Africa – Latin America and the Caribbean
	Country	See relevant country list
	Study design (broad)	<ul style="list-style-type: none"> – Randomized Controlled Trial (RCT) – Regression Discontinuity Design (RDD) – Controlled-Before and After Study (Treatment and Comparison Group, Baseline and Endline Data) – Panel data but no Baseline (Treatment and Comparison Group, Panel Data but no Baseline Data) – Comparison Group with Endline Data Only – Interrupted Time Series (ITS) – Other
	Study design / analysis	<ul style="list-style-type: none"> – Difference-in-Differences (DID) Estimation Methods – Fixed- or random effects model – Instrumental Variables (IV) – Propensity Score Matching (PSM) and other matching – Multivariate Regression – Comparison of Means – Other
Comments on study design	Open answer – any relevant information on the impact evaluation study design and analysis methods – for example if methods are	

Category		Answer
		combined or the authors use multiple methods separately.
	Mixed methods?	Yes/No This refers to whether or not the paper includes a qualitative component .
Intervention	Category of Intervention	– Refer to table 1
	Interventions	– Refer to table 1
	Intervention includes multiple PMM-components	Yes/No
	Intervention includes non-PMM component?	Yes/No
	Intervention(s) description	Open answer - This refers to the intervention description given by the review authors rather than the descriptions of interventions in each impact evaluation included in the EGM. This should be a succinct summary of the intervention – including a description of whether the intervention is a single intervention from the framework or a package of interventions.
Outcomes N.B. For the definitions, refer to the outcome definitions given by the review authors rather than the descriptions of outcomes in each primary study included in the review. Applies to each outcome definition question for SRs.	Outcomes	- Refer to table 2.
	Provider and managerial outputs and outcomes (effects at the level of individual providers and managerial staff)	Open answer, refers to the following outcomes: <ul style="list-style-type: none"> • Workload • Work morale • Stress • Burnout • Sick leave • Staff turnover • Knowledge • Attitudes, beliefs, perception • Skills and competencies
	Patient outcomes (defined as changes in health status and/or on patient health behaviors)	Open answer, refers to the following outcomes: <ul style="list-style-type: none"> • Health Status Outcomes: (1) Physical health • Health Status Outcomes: (2) Psychological health • Health Status: (3) Psychosocial health • Health Behaviors: (1) Adherence by patients • Health Behaviors: (2) Health seeking behaviors • Unintended patient outcomes

Category		Answer
	Organizational outcomes (defined as organizational-level within and across facilities and networks of primary care)	Open answer, refers to the following outcomes: <ul style="list-style-type: none"> • Quality of care process improvement; • Adherence to recommended practice or guidelines • Patient satisfaction • Perceived quality of care • Workforce retention • Changes in organizational culture • Unintended organizational outcomes (gaming, shirking, shaming, data falsification, etc.)
	Population health outputs and outcomes (defined as aggregate, health and equity effects accruing defined populations)	Open answer, refers to the following outcomes: <ul style="list-style-type: none"> • Utilization of specific primary care services; • Coverage of specific services or interventions (example: proportion of pregnant women receiving antenatal care; proportion of pregnant women delivering in facilities; coverage rate of specific vaccines); • Access to primary care services (example waiting times) • Health equity effects • Adverse health effects or harm • Unintended population health outcomes
	Social outcomes (defined as non-health, social, economic, or cultural effects affecting defined populations)	Open answer, refers to the following outcomes: <ul style="list-style-type: none"> • Community participation • Other equity effects • Other unintended social outcomes
Equity Data	How does this study consider equity?	<ul style="list-style-type: none"> • Does not address equity • Contains equity-sensitive analytical frameworks / theory of change • Uses equity-sensitive research questions • Follows equity-sensitive methodologies (sub-group analysis) • Contains equity-sensitive methodologies: additional study components to assess how and why (including mixed and qualitative methods) • Uses any other methodology that is equity sensitive that is not covered by the other options • Uses equity-informed research processes (who are the respondents, who collects data, when, where etc.)

Category		Answer
		<ul style="list-style-type: none"> • Addresses interventions targeting specific vulnerable groups - Looks at the impact of an intervention that targets specific population groups • Measures effects on an inequality outcome
	Dimension of equity/ Population group	<ul style="list-style-type: none"> • Age • Disability • Education • Gender, sex • Occupation • Other vulnerable groups • Place of residence • Race, ethnicity, culture and language • Religion • Social capital • Socioeconomic status
	Equity description	Open answer – provide a description of how the study considers equity, and for which group. This is to corroborate and elaborate on the answers above. For example, describe the sub-group analysis undertaken, how the intervention targets a disadvantaged group or how the authors used an equity sensitive framework to inform their study. Please also note the page number where this information can be found.
Access	Link	Open answer (if already on 3ie database, please use this link)
	On 3ie database (yet)?	Yes/No
	Publication status	<ul style="list-style-type: none"> - Journal Article - Published Working Paper - Book or Book Chapter - Conference Paper - Dissertation - Published Report - Unpublished Paper

B- Data extraction tool for systematic reviews (SR)

Category		Answer
Descriptive information	ID	Open answer - SRxxx
	Title	Open answer
	Full title	Open answer
	Author Citation	Open answer
	Publication date	Open answer
	Maps	Performance measurement and management in primary care systems in low and middle-income countries evidence gap map
	Regions (3ie)	<ul style="list-style-type: none"> - East Asia and Pacific - South Asia - Europe - CIS - Middle East and North Africa - Sub-Saharan Africa - Latin America and the Caribbean - North America - Global - Unclear
	SR includes studies from high-income countries?	Yes/No
	Number of impact evaluations included (Systematic review only)	Open answer
Intervention	Category of Intervention	- Refer to Table 1
Intervention/ outcome	Interventions	- Refer to Table 1
	Intervention(s) description	Open answer. -
	SR covers non-PMM interventions	Yes/No
	Outcomes	Refer to Table 2 in the protocol
	Provider and managerial outputs and outcomes (effects at the level of individual providers and managerial staff)	Open answer, refers to the following outcomes: <ul style="list-style-type: none"> • Workload • Work morale • Stress • Burnout • Sick leave • Staff turnover • Knowledge • Attitudes, beliefs, perception
<i>N.B. This refers to the intervention description given by the review authors rather than the descriptions of interventions in each primary study included in the review.</i>		

Category		Answer
		<ul style="list-style-type: none"> • Skills and competencies
Outcome definitions	Patient outcomes (defined as changes in health status and/or on patient health behaviors)	<p>Open answer, refers to the following outcomes:</p> <ul style="list-style-type: none"> • Health Status Outcomes: (1) Physical health • Health Status Outcomes: (2) Psychological health • Health Status: (3) Psychosocial health • Health Behaviors: (1) Adherence by patients • Health Behaviors: (2) Health seeking behaviors • Unintended patient outcomes
	Organizational outcomes (defined as organizational-level within and across facilities and networks of primary care)	<p>Open answer, refers to the following outcomes:</p> <ul style="list-style-type: none"> • Quality of care process improvement; • Adherence to recommended practice or guidelines • Patient satisfaction • Perceived quality of care • Workforce retention • Changes in organizational culture <p>Unintended organizational outcomes (gaming, shirking, shaming, data falsification, etc.)</p>
	Population health outputs and outcomes (defined as aggregate, health and equity effects accruing defined populations)	<p>Open answer, refers to the following outcomes:</p> <ul style="list-style-type: none"> • Utilization of specific primary care services; • Coverage of specific services or interventions (example: proportion of pregnant women receiving antenatal care; proportion of pregnant women delivering in facilities; coverage rate of specific vaccines); • Access to primary care services (example waiting times) • Health equity effects • Adverse health effects or harm • Unintended population health outcomes

Category		Answer
	Social outcomes (defined as non-health, social, economic, or cultural effects affecting defined populations)	Open answer, refers to the following outcomes: <ul style="list-style-type: none"> • Community participation • Other equity effects • Other unintended social outcomes •
	How does this study consider equity?	<ul style="list-style-type: none"> • Does not address equity • Contains equity-sensitive analytical frameworks / theory of change • Uses equity-sensitive research questions • Follows equity-sensitive methodologies (sub-group analysis) • Contains equity-sensitive methodologies: additional study components to assess how and why (including mixed and qualitative methods) • Uses any other methodology that is equity sensitive that is not covered by the other options • Uses equity-informed research processes (who are the respondents, who collects data, when, where etc.) • Addresses interventions targeting specific vulnerable groups - Looks at the impact of an intervention that targets specific population groups • Measures effects on an inequality outcome
Equity Data	Dimension of equity/ Population group	Open answer, refers to the following outcomes: <ul style="list-style-type: none"> • Other non-health equity effects • Adverse non-health effects or harm • Community participation • Other unintended social outcomes
	Dimension of equity/ Population description	<ul style="list-style-type: none"> • Place of residence • Ethnicity, culture and language • Gender • Religion

Category		Answer
		<ul style="list-style-type: none"> • Education • Socioeconomic status • Social capital • Age • Disability • Other (vulnerable group not typified by any of the above)
	Equity description	Open answer – provide a description of how the study considers equity, and for which group. This is to corroborate and elaborate on the answers above. For example, describe the sub-group analysis undertaken, how the intervention targets a disadvantaged group or how the authors used an equity sensitive framework to inform their study. Please also note the page number where this information can be found.
Review confidence	Confidence in review (taken from quality appraisal from the adapted version of the SURE checklist) - (Systematic review only)	<ul style="list-style-type: none"> – High – Medium – Low
	If high confidence, summary of findings - (Systematic review only)	Open answer
Access	Link	Open answer (if already on 3ie database, please use this link)
	On 3ie database (yet)?	Yes/No

Data to be collected and summarized for database entries: terms of reference

- (1) Existing record on database: YES/ NO (Check if there is an existing record on the database, if yes, work from this and update it, if no proceed making a new record)
- (2) Title:
- (3) Author:
- (4) Geographical coverage: Please list with standard regions used for the database. These are: *East Asia and Pacific (including South East Asia), South Asia, Europe and CIS, Middle East and North Africa, Sub-Saharan Africa, Latin America and the Caribbean, North America, Commonwealth of Independent States, Developed Countries*
- (5) Sector: Please use list with standard sectors for the database. More than one sector category, or more sub-sectors within this category, may apply. If in doubt, err on the side of being over-inclusive, and leave a comment for the peer reviewer indicating any doubts. These are: *reproductive, maternal, newborn, child or adolescent*
- (6) Sub-sector: See notes at the bottom of the document for full list of sub-sectors*
- (7) Equity focus: this might be no, or one/ several of the following: *gender, indigenous groups, ethnic minorities, differently-abled, orphans and vulnerable children, elderly, migrant workers, refugees*. If equity considered, please state in what way.
- (8) Status: Completed, Protocol or Title stage (delete as appropriate)
- (9) Review type: effectiveness review (drawing on evidence from impact evaluations), or other review (delete as appropriate)
- (10) Quantitative synthesis method, if applicable: e.g.: narrative/thematic synthesis; meta-analysis etc?
- (11) Qualitative synthesis method, if applicable:
- (12) For systematic reviews rated as high-confidence only:
 - (13) Background: Brief description of the interventions and motivation for the review. What is the problem? What is the intervention? How does it aim to have an impact on outcomes of interest?
 - (14) Objectives: Objectives of the review – here you can quote authors verbatim, using quotation marks. But the authors stated objectives are usually a bit too long/wordy and/or any one statement does not always contain all of the objectives (this information might be at several places in the review). Rewording is almost always possible and the resulting reworded statement is usually more concise and more easily understandable, so when please reword the objectives when possible.
 - (15) Main findings: This section should include a brief description of findings, including number of included studies, location of studies, results of synthesis (if meta-analysis was conducted, include pooled effect size and 95 percent confidence interval for relevant syntheses), conclusions, findings regarding methodology and future research.

The section should be structured as follows (including the subheadings):

One line summary: summarise the conclusions of the review in one or two sentences

Evidence base: Number and types of studies, geographical location and thematic focus

Policy relevant findings: findings related to the effectiveness of the intervention etc

Implications for further research: any mention of issues to be addressed in future research, including issues relating to study design.

- (16) Methodology: Inclusion criteria (including population, intervention, study design (e.g.: RCTs and Quasi-experimental studies), outcomes and contexts), outline of search (including main databases and time period of search, data collection and synthesis. Please see guidelines for more guidance on this section.
- (17) Applicability/external validity: Does the review discuss how generalizable the results are? What methods, if any, does the review use to assess applicability/external validity? Do the authors take any steps to improve the applicability/external validity of the findings of the review (eg: use a theory-based approach, drawing on a logic model or program theory, and/ or reporting information along the causal chain)
- (18) Publication Source: Author, year, Title, Publication details (if journal: title, vol, no, pp; if report series/ working paper etc.: Title of series and no if applicable, publisher location: publisher name; if book: publisher location: publisher name)
- Example from JDEFF citation for journal article citation (for more refer JDEFF)
- Salazar, D.J. and Alper, D.K. (2002) Reconciling environmentalism and the left: perspectives on democracy and social justice in British Columbia's environmental movement. *Canadian Journal of Political Science*, 35(4), pp. 527–566.
- (19) Downloadable link: Include link to downloadable paper – preferably open access, but if this is not available include link to pay per view.
- (20) Contact details of corresponding author/s:
- (21) Summary of quality assessment (from C3 in checklist below):

Supplementary file 2 - Part 2 - Systematic reviews critical appraisal checklist (adapted version of SURE checklist)ⁱ

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?</p> <p>Did the authors specify:</p> <p><input type="checkbox"/> Types of studies</p> <p><input type="checkbox"/> Participants/ settings/ population</p> <p><input type="checkbox"/> Intervention(s)</p> <p><input type="checkbox"/> Outcome(s)</p>	<p><input type="checkbox"/> Yes</p> <p><input type="checkbox"/> Partially</p> <p><input type="checkbox"/> No</p> <p><i>Coding guide - check the answers above</i></p> <p><i>YES: All four should be yes</i></p> <p><i>NO: All four should be no</i></p> <p><i>PARTIALLY: Any other</i></p>
<p><i>Comments (note important limitations or uncertainty)</i></p>	

<p>A.2 Was the search for evidence reasonably comprehensive?</p> <p>Were the following done:</p> <p><input type="checkbox"/> Language bias avoided (no restriction of inclusion based on language)</p> <p><input type="checkbox"/> No restriction of inclusion based on publication status</p> <p><input type="checkbox"/> Relevant databases searched (<u>Minimum criteria:</u> All reviews should search at least one source of grey literature such as Google; for health: Medline/ Pubmed + Cochrane Library; for social sciences IDEAS + at least one database of general social science literature and one subject specific database)</p> <p><input type="checkbox"/> Reference lists in included articles checked</p> <p><input type="checkbox"/> Authors/experts contacted</p>	<p><input type="checkbox"/> Yes</p> <p><input type="checkbox"/> Partially</p> <p><input type="checkbox"/> No</p> <p><input type="checkbox"/> Can't tell</p> <p><i>Coding guide - check the answers above:</i></p> <p><i>YES: All five should be yes</i></p> <p><i>PARTIALLY: Relevant databases and reference lists are both reported</i></p> <p><i>NO: Any other</i></p>
<p><i>Comments (note important limitations or uncertainty)</i></p>	

<p>A.3 Does the review cover an appropriate time period?</p> <p><i>Is the search period comprehensive enough that relevant literature is unlikely to be omitted?</i></p>	<p><input type="checkbox"/> Yes</p> <p><input type="checkbox"/> Can't tell (only use if no information about time period for search)</p> <p><input type="checkbox"/> No</p> <p><input type="checkbox"/> Unsure</p> <p><i>Coding guide:</i></p> <p><i>YES: Generally, this means searching the literature at least back to 1990</i></p> <p><i>NO: Generally, if the search does not go back to 1990</i></p> <p><i>CAN'T TELL: No information about time period for search</i></p> <p><i>Note: With reference to the above – there may be important reasons for adopting different dates for the search, e.g. depending on the intervention. If you think there are limitations with the timeframe adopted for the search which have not been noted and justified by the authors, you should code this item as a NO and specify your reason for doing so in the comment box below. Older reviews should not be downgraded, but the fact that the search was conducted some time ago should be noted in the quality assessment. Always report the time period for the search in the comment box.</i></p>
<p><i>Comments (note search period, any justification provided for the search period, or uncertainty)</i></p>	

<p>A.4 Was bias in the selection of articles avoided?</p> <p>Did the authors specify:</p> <p><input type="checkbox"/> Independent screening of full text by at least 2 reviewers</p> <p><input type="checkbox"/> List of included studies provided</p> <p><input type="checkbox"/> List of excluded studies provided</p>	<p><input type="checkbox"/> Yes</p> <p><input type="checkbox"/> Partially</p> <p><input type="checkbox"/> No</p> <p><i>Coding guide:</i></p> <p><i>YES: All three should be yes, although reviews published in journals are unlikely to have a list of excluded studies (due to limits on word count) and the review should not be penalised for this.</i></p> <p><i>PARTIALLY: Independent screening and list of included studies provided are both reported</i></p> <p><i>NO: All other. If list of included studies provided, but the authors do not report whether or not the screening has been done by 2 reviewers review is downgraded to NO.</i></p>
<p><i>Comments (note important limitations or uncertainty):</i></p>	
<p>A.5 Did the authors use appropriate criteria to assess the quality and risk of bias in analyzing the studies that are included?ⁱⁱ</p> <p><input type="checkbox"/> The criteria used for assessing the quality/ risk of bias were reported</p> <p><input type="checkbox"/> A table or summary of the assessment of each included study for each criterion was reported</p> <p><input type="checkbox"/> Sensible criteria were used that focus on the quality/ risk of bias (and not other qualities of the studies, such as precision or applicability/external validity). “Sensible” is defined as a recognized quality appraisal tool/ checklist, or similar tool which assesses bias in included studies. Please see footnotes for details of the main types of bias such a tool should assess.</p>	<p><input type="checkbox"/> Yes</p> <p><input type="checkbox"/> Partially</p> <p><input type="checkbox"/> No</p> <p><i>Coding guide:</i></p> <p><i>YES: All three should be yes</i></p> <p><i>PARTIALLY: The first and third criteria should be reported. If the authors report the criteria for assessing risk of bias and report a summary of this assessment for each criterion, but the criteria may be only partially sensible (e.g. do not address all possible risks of bias, but do address some), we downgrade to PARTIALLY.</i></p> <p><i>NO: Any other</i></p>

<i>Comments (note important limitations or uncertainty)</i>	
<p>A.6 Overall – how much confidence do you have in the methods used to identify, include and critically appraise studies?</p> <p><i>Summary assessment score A relates to the 5 questions above.</i></p> <p><i>High confidence applicable when the answers to the questions in section A are all assessed as ‘yes’</i></p> <p><i>Low confidence applicable when any of the following are assessed as ‘NO’ above: not reporting explicit selection criteria (A1), not conducting reasonably comprehensive search (A2), not avoiding bias in selection of articles (A4), not assessing the risk of bias in included studies (A5)</i></p> <p><i>Medium confidence applicable for any other – i.e. section A3 is assessed as ‘NO’ or ‘can’t tell’ and remaining sections are assessed as ‘partially’ or ‘can’t tell’</i></p>	<p><input type="checkbox"/> Low confidence (limitations are important enough that the results of the review are not reliable)</p> <p><input type="checkbox"/> Medium confidence (limitations 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"/> High confidence (only minor limitations)</p>
<i>Comments (note important limitations).</i>	

Section B: Methods used to analyze the findings

<p>B.1 Were the characteristics and results of the included studies reliably reported?</p> <p>Was there:</p> <p><input type="checkbox"/> Independent data extraction by at least 2 reviewers</p> <p><input type="checkbox"/> A table or summary of the characteristics of the participants, interventions and outcomes for the included studies</p> <p><input type="checkbox"/> A table or summary of the results of all the included studies</p>	<p><input type="checkbox"/> Yes</p> <p><input type="checkbox"/> No</p> <p><input type="checkbox"/> Partially</p> <p><input type="checkbox"/> Not applicable (e.g. no included studies)</p> <p><i>Coding guide:</i></p> <p><i>YES: All three should be yes</i></p> <p><i>PARTIALLY: Criteria one and three are yes, but some information is lacking on second criteria.</i></p> <p><i>No: None of these are reported. If the review does not report whether data was independently extracted by 2 reviewers (possibly a reporting error), we downgrade to NO.</i></p> <p><i>NOT APPLICABLE: if no studies/no data</i></p>
<p><i>Comments (note important limitations or uncertainty)</i></p>	
<p>B.2 Are the methods used by the review authors to analyze the findings of the included studies clear, including methods for calculating effect sizes if applicable?</p>	<p><input type="checkbox"/> Yes</p> <p><input type="checkbox"/> Partially</p> <p><input type="checkbox"/> No</p> <p><input type="checkbox"/> Not applicable (e.g. no studies or no data)</p> <p><i>Coding guide:</i></p> <p><i>YES: Methods used clearly reported. If it is clear that the authors use narrative synthesis, they don't need to say this explicitly.</i></p> <p><i>PARTIALLY: Some reporting on methods but lack of clarity</i></p> <p><i>NO: Nothing reported on methods</i></p> <p><i>NOT APPLICABLE: if no studies/no data</i></p>
<p><i>Comments (note important limitations or uncertainty)</i></p>	

<p>B.3 Did the review describe the extent of heterogeneity?</p> <p><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?</p> <p><input type="checkbox"/> Did the review discuss the extent to which there were important differences in the results of the included studies?</p> <p><input type="checkbox"/> If a meta-analysis was done, was the I^2, chi square test for heterogeneity or other appropriate statistic reported? If no statistical test was reported, is a qualitative justification made for the use of random effects?</p>	<p><input type="checkbox"/> Yes</p> <p><input type="checkbox"/> Partially</p> <p><input type="checkbox"/> No</p> <p><input type="checkbox"/> Not applicable (e.g. no studies or no data)</p> <p><i>Coding guide:</i></p> <p><i>YES: First two should be yes, and third category should be yes if applicable should be yes</i></p> <p><i>PARTIALLY: The first category is yes</i></p> <p><i>NO: Any other</i></p> <p><i>NOT APPLICABLE: if no studies/no data</i></p>
<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 (sample size) <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 	<ul style="list-style-type: none"> <input type="checkbox"/> Yes <input type="checkbox"/> Partially <input type="checkbox"/> No <input type="checkbox"/> Not applicable (e.g. no studies or no data) <input type="checkbox"/> Can't tell <p><i>Coding guide:</i></p> <p><i>YES: If appropriate table, graph or meta-analysis AND appropriate weights AND unit of analysis errors addressed (if appropriate).</i></p> <p><i>PARTIALLY: If appropriate table, graph or meta-analysis AND appropriate weights AND unit of analysis errors not addressed (and should have been).</i></p> <p><i>NO: If narrative OR vote counting (where quantitative analyses would have been possible) OR inappropriate reporting of table, graph or meta-analyses.</i></p> <p><i>NOT APPLICABLE: if no studies/no data</i></p> <p><i>CAN'T TELL: if unsure (note reasons in comments below)</i></p>
<p><i>Comments (note important limitations or uncertainty)</i></p>	

<p>B. 5 Does the review report evidence appropriately?</p> <p><input type="checkbox"/> The review makes clear which evidence is subject to low risk of bias in assessing causality (attribution of outcomes to intervention), and which is likely to be biased, and does so appropriately</p> <p><input type="checkbox"/> Where studies of differing risk of bias are included, results are reported and analyzed separately by risk of bias status</p>	<p><input type="checkbox"/> Yes</p> <p><input type="checkbox"/> No</p> <p><input type="checkbox"/> Partially</p> <p><input type="checkbox"/> Not applicable</p> <p><i>Coding guide:</i></p> <p><i>YES: Both criteria should be fulfilled (where applicable)</i></p> <p><i>NO: Criteria not fulfilled</i></p> <p><i>PARTIALLY: Only one criteria fulfilled, or when there is limited reporting of quality appraisal (the latter applies only when inclusion criteria for study design are appropriate)</i></p> <p><i>NOT APPLICABLE: No included studies</i></p> <p><i>Note on reporting evidence and risk of bias: For reviews of effects of 'large n' interventions, experimental and quasi-experimental designs should be included (if available). For reviews of effects of 'small n' interventions, designs appropriate to attribute changes to the intervention should be included (e.g. pre-post with assessment of confounders)</i></p>
<p><i>Please specify included study designs and any other comments (note important limitations or uncertainty):</i></p>	

<p>B.6 Did the review examine the extent to which specific factors might explain differences in the results of the included studies?</p> <p><input type="checkbox"/> Were factors that the review authors considered as likely explanatory factors clearly described?</p> <p><input type="checkbox"/> Was a sensible method used to explore the extent to which key factors explained heterogeneity?</p> <ul style="list-style-type: none"> <input type="checkbox"/> Descriptive/textual <input type="checkbox"/> Graphical <input type="checkbox"/> Meta-analysis by sub-groups <input type="checkbox"/> Meta-regression <input type="checkbox"/> Other 	<p><input type="checkbox"/> Yes</p> <p><input type="checkbox"/> Partially</p> <p><input type="checkbox"/> No</p> <p><input type="checkbox"/> Not applicable</p> <p><i>Coding guide:</i></p> <p><i>YES: Explanatory factors clearly described and appropriate methods used to explore heterogeneity</i></p> <p><i>PARTIALLY: Explanatory factors described but for meta-analyses, sub-group analysis or meta-regression not reported (when they should have been)</i></p> <p><i>NO: No description or analysis of likely explanatory factors</i></p> <p><i>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</i></p>
<p><i>Comments (note important limitations or uncertainty)</i></p>	

<p>B.7 Overall - how much confidence do you have in the methods used to analyze 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>High confidence applicable when all the answers to the questions in section B are assessed as 'yes'.</i></p> <p><i>Low confidence applicable when any of the following are assessed as 'NO' above: critical characteristics of the included studies not reported (B1), not describing the extent of heterogeneity (B3), combining results inappropriately (B4), reporting evidence inappropriately (B5).</i></p> <p><i>Medium confidence applicable for any other: i.e. the "Partial" option is used for any of the 6 preceding questions or questions and/or B.2 and/ or B.6 are assessed as 'no'.</i></p>	<p><input type="checkbox"/> Low confidence (limitations are important enough that the results of the review are not reliable)</p> <p><input type="checkbox"/> Medium confidence (limitations 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"/> High confidence (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 – only one person reviewing</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 Are there any mitigating factors which should be taken into account in determining the reviews reliability?</p>	<p><input type="checkbox"/> Limitations acknowledged</p> <p><input type="checkbox"/> No strong policy conclusions drawn (including in abstract/ summary)</p> <p><input type="checkbox"/> Any other factors</p>
<p><i>Use comments to specify if relevant, to flag uncertainty or need for discussion</i></p>	
<p>C.3 Based on the above assessments of the methods how would you rate the reliability of the review?</p> <p><input type="checkbox"/> <u>Low confidence in conclusions about effects:</u> The systematic review has the following major limitations...</p> <p><input type="checkbox"/> <u>Medium confidence in conclusions about effects:</u> The systematic review has the following limitations...</p> <p><input type="checkbox"/> <u>High confidence in conclusions about effects:</u> If applicable: The review has the following minor limitations...</p> <p><i>Coding guide:</i></p> <p>High confidence in conclusions about effects: high confidence noted overall for sections A and B, unless moderated by answer to C1.</p> <p>Medium confidence in conclusions about effects: medium confidence noted overall for sections A or B, unless moderated by answer to C1 or C2.</p> <p>Low confidence in conclusions about effects: low confidence noted overall for sections A or B, unless moderated by answer to C1 or C2.</p>	

Limitations should be summarized above, based on what was noted in Sections A, B and C.

NOTES

ⁱ Adapted from Supporting the Use of Research Evidence (SURE) Collaboration. SURE checklist for making judgements about how much confidence to place in a systematic review. In: SURE guides for preparing and using policy briefs. www.evipnet.org/sure

ⁱⁱ **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 social, economic and health care interventions, the main types of bias arise from systematic differences in the groups that are compared (selection bias), the intervention that is provided, or exposure to other factors apart from the intervention of interest (performance bias/contamination), withdrawals or exclusions of people entered into a study (attrition bias) or how outcomes are assessed (detection bias) and reported (reporting bias). Reviews of social science studies may be particularly affected by reporting bias, where a biased subset of all the relevant data and analyses is presented.

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.

***Question 6: Full list of sub-sectors by sector**

Agriculture and Rural Development

Agricultural Reform; Agricultural Credit; Agricultural Extension; Agro-Industry & Marketing; Weather Insurance; Fisheries & Aquaculture; Forestry; Irrigation & Drainage; Livestock; Agricultural Research; Rural Land Reform; Rural Livelihoods; Rural Roads

Economic Policy

Macro/ Non-Trade; Trade; Tax Reform

Education

Distance Education/ Education Technology; Educational Inputs; Girl's Education; Non-formal Education; Pre-Primary and Primary Education; Public/ Private Sector Education; Secondary Education; Student Loan; System Reform & Capacity Building; Tertiary Education; Vocational/ Technical Education & Training

Energy

Distribution & Transmission; Hydro; Thermal; Other Power & Energy Conversion; Rural Electrification

Environment and Disaster Management

Biodiversity; Climate Change; Environmental Institutions; Natural Resource Management; Pollution Control/ Waste Management; Resettlement; Disaster Relief

Finance

Capital Markets Development; Banking Systems; Consumer Credit; Financial Sector Reform; Microfinance;

Health Nutrition and Population

Child Nutrition; Nutrition; Health Sector Reform; Health Financing, Insurance and User Fees; Health Services; HIV/AIDS; Hospitals- Secondary & Tertiary; Mortality; Preventive Health and Health Behavior; Primary Health- including reproductive health; Sexual Behavior; Specific Diseases- including Malaria TB; Specific Diseases-diarrhoea.

Information and Communication Technology

Telecommunications Mass Media Technological Innovation Technology Development Funds

Multisector

Community Action Program; Community Driven Development; Concessions; Conditional Cash Transfers; Conflict Prevention and Post-Conflict Reconstruction; Social Funds; Early Childhood Development Programs; Multisector

Private Sector Development

Business Environment; Private Infrastructure; Small Scale Enterprise

Public Sector Management

Anti-corruption/ Governance; Civil Service Reform; Decentralization; Institutional Development; Judicial Reform; Public Financial Management

Social Protection

Labor markets & Employment; Pensions & Social Insurance; Savings and Remittances; Social Assistance; Social Protection Reform

Transportation

Highways; Ports and Waterways Railways

Urban Development

Urban housing; Urban land reform; Urban transport; Slum upgradation Programmes; Urban development and management

Water and Sanitation

Water Supply and Sanitation Reform; Urban Water and Sanitation; Rural Water and Sanitation