Open access

Original research

Monica Jain ‍ ‍,1 Shannon Shisler,2 Charlotte Lane ‍ ‍,3 Avantika Bagai,4
Elizabeth Brown,5 Mark Engelbert ‍ ‍3

To cite: Jain M, Shisler S,
Lane C, et al. Use of community
engagement interventions to
improve child immunisation in
low-­income and middle-­income
countries: a systematic review
and meta-­analysis. BMJ Open
2022;12:e061568. doi:10.1136/
bmjopen-2022-061568
► Prepublication history and
additional supplemental material
for this paper are available
online. To view these files,
please visit the journal online
(http://dx.doi.org/10.1136/​
bmjopen-2022-061568).

Received 09 February 2022
Accepted 10 October 2022

© Author(s) (or their
employer(s)) 2022. Re-­use
permitted under CC BY.
Published by BMJ.
1

International Initiative for
Impact Evaluation, New Delhi,
India
2
International Initiative for
Impact Evaluation, London, UK
3
International Initiative for
Impact Evaluation, Washington,
District of Columbia, USA
4
Development Solutions, New
Delhi, India
5
Center for Effective Global
Action, University of California,
Berkeley, California, USA
Correspondence to
Dr Monica Jain;
m
​ jain@​3ieimpact.o​ rg

ABSTRACT
Objective To support evidence informed decision-­making,
we systematically examine the effectiveness and cost-­
effectiveness of community engagement interventions
on routine childhood immunisation outcomes in low-­
income and middle-­income countries (LMICs) and identify
contextual, design and implementation features associated
with effectiveness.
Design Mixed-­methods systematic review and meta-­
analysis.
Data sources 21 databases of academic and grey
literature and 12 additional websites were searched in
Eligibility criteria for selecting studies We included
experimental and quasi-­experimental impact evaluations
of community engagement interventions considering
outcomes related to routine child immunisation in LMICs.
No language, publication type, or date restrictions were
imposed.
Data extraction and synthesis Two independent
researchers extracted summary data from published
reports and appraised quantitative risk of bias using
adapted Cochrane tools. Random effects meta-­analysis
was used to examine effects on the primary outcome, full
immunisation coverage.
Results Our search identified over 43 000 studies
and 61 were eligible for analysis. The average pooled
effect of community engagement interventions on full
immunisation coverage was standardised mean difference
0.14 (95% CI 0.06 to 0.23, I2=94.46). The most common
source of risk to the quality of evidence (risk of bias) was
outcome reporting bias: most studies used caregiver-­
reported measures of vaccinations received by a child in
the absence or incompleteness of immunisation cards.
Reasons consistently cited for intervention success include
appropriate intervention design, including building in
community engagement features; addressing common
contextual barriers of immunisation and leveraging
facilitators; and accounting for existing implementation
constraints. The median intervention cost per treated
child per vaccine dose (excluding the cost of vaccines) to
increase absolute immunisation coverage by one percent
was US$3.68.

STRENGTHS AND LIMITATIONS OF THIS STUDY
⇒ Thorough literature search of 21 major electronic

⇒

⇒

⇒

⇒

databases and reporting as per Preferred Reporting
Items for Systematic Reviews and Meta-­Analyses
guidelines.
Presents a nuanced framework of community engagement with a typology that differentiates three
types of interventions: interventions in which community engagement is embedded, and those interventions that engage community in their design or
implementation.
The effects of community engagement interventions
are robust to exclusion of studies assessed as high
risk of bias for almost all the primary outcomes. The
effects are also uniform across geographies and
baseline immunisation rates.
For some immunisation outcomes the evidence
base for drawing conclusions is adequate, for others
it is limited.
Evidence base is skewed across the three engagement types with a relatively large evidence
base for those interventions in which community
engagement is embedded and limited for those
interventions using engagement in implementation
autonomy

Conclusion Community engagement interventions are
successful in improving outcomes related to routine child
immunisation. The findings are robust to exclusion of
studies assessed as high risk of bias.

INTRODUCTION
Immunisation is one of the most cost-­effective
ways to prevent and control life-­threatening
infectious diseases. From 2001 to 2020, projects that introduced or increased coverage
of vaccines averted an estimated 14 million
deaths, 350 million cases of illness, 8 million
cases of long-­term disability and 700 million
disability-­adjusted life-­years.1 Nonetheless,


1

BMJ Open: first published as 10.1136/bmjopen-2022-061568 on 8 November 2022. Downloaded from http://bmjopen.bmj.com/ on October 21, 2023 by guest. Protected by copyright.

Use of community engagement
interventions to improve child
immunisation in low-­income and middle-­
income countries: a systematic review
and meta-­analysis


rates of routine vaccination of children in low-income and middle-income countries (LMICs) are low or stagnant. In 2019, an estimated 19.7 million infants did not receive routine immunisations. Around 60% of these children live in ten LMICs, including Ethiopia, India, Nigeria and Pakistan as of 2019.2

Community engagement approaches feature prominently in global immunisation strategies.3 However, there is a dearth of rigorous and systematic evidence on effectiveness of community engagement approaches to improve routine childhood immunisations specifically in LMICs. In our search, we could find only two systematic reviews for LMICs, which analysed effectiveness of community monitoring interventions and preventive interventions delivered by community health workers, respectively.45 As such, previous systematic reviews do not provide adequate guidance to stakeholders interested in understanding whether and how alternative community engagement interventions work in LMICs to improve routine childhood immunisations at what cost. There is, therefore, a need to make such evidence available to guide policymakers and public health practitioners in making informed decisions about these interventions. To address this knowledge gap, we conducted a systematic review examining the effects of community engagement interventions on outcomes related to childhood immunisation in LMICs, determining their cost-effectiveness and identifying contextual, design and implementation features that may be associated with intervention effectiveness.

METHODS

Overview
The protocol of this systematic review with meta-analysis is registered with The Campbell Collaboration.6 We followed the Campbell and Cochrane Collaborations’ guidelines for systematic reviewing7 and drew on theory-based mixed-methods impact evaluation8 and systematic review12 concepts. We followed the PRISMA reporting guidelines. The amendments to the information provided in the protocol is reported in online supplemental appendix 1.

Conceptual framework
For our review, we defined ‘communities’ in reference to the lowest level of the health service delivery system (or whatever level provides routine immunisation services in the local context). A community is a group of people who serve or are served by a particular primary health facility. Thus, communities encompass a wide range of stakeholders, including caregivers, health service providers and influential community members such as religious or other traditional leaders.

WHO 2020 defines community engagement as ‘a process of developing relationships that enable stakeholders to work together to address health-related issues and promote well-being to achieve positive health impact and outcomes’.14 For this review, we developed a framework that classified community engagement interventions based on process of engagement as in the WHO definition. It also corresponds to the ‘utilitarian perspective’ of community engagement captured and articulated in Brunton et al15. ‘In utilitarian perspectives, health (and other) services reach out to engage particular communities that they have identified require assistance and the intervention is devised within existing policy, practice and resource frameworks.’ In addition, our framework goes beyond one-way communication to include some consultation or dialogue with the community or some decision-making by them. We considered three points within an intervention during which engagement could occur, as elaborated below and in online supplemental appendix 2.

Engagement in the design of interventions: Community input or feedback was sought before implementing an intervention (eg, pilot, needs assessment, formative evaluation and outreach).

Engagement in implementation autonomy of interventions: Community was used in intervention implementation as healthcare workers, facilitators or problem solvers and only if they had some opportunity to affect or influence its implementation.

Engagement as the intervention (engagement is embedded): A serious attempt was made to gain community buy-in for activities or new cadres of community-based structures were established (eg, village health committees or community health volunteers).

Research questions
The research questions for this review were:
1. What evidence exists regarding the effectiveness of community engagement interventions in improving routine immunisation coverage of children in LMICs?
2. Is there evidence for heterogeneous effects of community engagement strategies (ie, does effectiveness vary by geographical region, gender or programme implementation)?
3. What factors relating to programme design, implementation and context are associated with better or worse outcomes along the causal chain? Do these vary by the kind of community engagement?
4. What is the cost-effectiveness of different community engagement interventions in improving children routine immunisation outcomes?

Search strategy
We implemented a systematic and comprehensive search strategy, in consultation with an information specialist. In May 2019 and May 2020, we searched 17 academic databases for experimental and quasi-experimental impact evaluations of community engagement interventions considering outcomes related to routine child immunisation in LMICs (using the World Bank country income classifications to determine LMIC status at the time the intervention began). We also searched 17 additional websites for grey literature. The list of sources searched and an example set of search strings are provided in
online supplemental appendix 3. We complemented this with citation tracking and contacting experts. The grey literature search was conducted by AB with support from external consultant reviewers. Given the limitations of the search functions on websites we searched for grey literature, it was not possible to use the same complex search strings used in academic databases, and search strategies were developed on a site-by-site basis.

**Inclusion/exclusion criteria (population, intervention, comparators, outcomes and study designs)**

The population, intervention, comparators, outcomes and study designs eligible for inclusion in the study are provided in table 1. No language, publication type or date restrictions were imposed. Because our definition of community focused on the lowest levels of health facilities, we excluded interventions targeting higher levels of the health system (eg, state-level officials) (online supplemental appendix 4). The primary outcomes considered in this review were coverage rates for (A) full immunisation, which is typically defined as the percentage of 1 year old who have received one dose of Bacille Calmette-Guérin (In some countries, other vaccinations such as those for JE encephalitis and yellow fever are administered to children as a part of the routine immunisation schedule. In those contexts, we went by the definition of full immunisation mentioned in the impact evaluation study), (B) third dose of DPT or pentavalent, (C) first dose of measles or (D) the timeliness of any of these doses. Additional antigen-specific immunisation coverage outcomes and secondary outcomes reflecting upstream conditions (eg, attitudes about vaccination and access to immunisation services) and downstream effects (eg, morbidity and mortality) of the primary outcomes were also included. Official health records and parent recall were considered acceptable sources of measures of immunisation coverage. The former was used when both measures were reported separately.

This review includes experimental and quasi-experimental studies that estimate the causal impact of an intervention, as compared with usual practice, by establishing a counterfactual. Specifically, studies with the following evaluation designs are included: randomised controlled trials, regression discontinuity designs, instrumental variables’ estimation, statistical matching (eg, propensity score matching), difference-in-differences (or any mathematical equivalent), fixed effects estimation and interrupted time series. We excluded studies for which the reported quantitative data could not be meaningfully converted to an effect size. In cases of relevant missing or incomplete data, we contacted study authors to obtain the required information. If we were unable to obtain the necessary data, we reported the characteristics of the study but did not include these studies in the meta-analysis. We conducted additional searches for economic and qualitative evidence on the included impact evaluations (online supplemental appendix 5).

**Screening**

At both the title and abstract and full-text screening stages, all papers were double screened by research consultants and supervised by MJ, ME and AB. Reconciliation meetings were held to resolve disagreements, and MJ and ME made final decisions on unresolved cases. The same reviewers manually searched for qualitative papers and project documents on Google Scholar and websites of implementing organisations and screened the papers for inclusion as they were identified.

**Data analysis**

Studies were coded for their engagement type by two reviewers (MJ and AB) who independently reviewed the intervention description and coded these against the definitions provided above. If studies allowed for engagement at several stages of the intervention, they could be coded as having more than one engagement type. We used Microsoft Excel to extract descriptive information and effect sizes from included studies using double coding. Coders reconciled their answers, and a study author made final decisions in case of disagreements. For qualitative analysis, all impact evaluations and additional documentation identified in the search were coded in NVivo. Cost data were single coded and checked by a study author (online supplemental appendix 6).

To avoid double-counting of evidence from different papers focusing on the same study, we linked these papers prior to analysis. We extracted data from the most recent publication. When data were reported over multiple time
periods, we extracted data for each period. Where authors reported the same outcome using more than one analytical model, we extracted data from the authors’ preferred model specification. When the preference was not specified, we used the model with the most controls. Where studies reported outcomes related to multiple treatment arms and only one comparison group, we estimated an effect size for each of the treatment arms.

To assess quantitative risk of bias, we created an adapted version of the Cochrane guidelines for assessing randomised controlled trials and non-randomised studies. In addition, we carried out a sensitivity analysis in which we considered only the high-quality qualitative studies that had a risk of bias assessment score of 20 or higher, indicating low risk of bias. For cost evidence, we assessed risk of bias along six primary dimensions adapted from a combination of tools, including: Doocy and Tappis; Campbell Collaboration Economic Methods Policy Brief and Methods for the Economic Evaluation of Health Care Programmes (online supplemental appendix 7).

We calculated the standardised mean difference, or Cohen’s d, its variance and SE for each effect, converting effects reported in other metrics as necessary, using formulae provided in Borenstein et al. In all cases we then adjusted Cohen’s d to Hedges’ g as defined in Ellis. For studies reporting regression results, we followed the approach of Keef and Roberts using the regression coefficient and the pooled SD of the outcome.

The amount of heterogeneity (ie, $\tau^2$) was estimated using the DerSimonian-Laird estimator. The Q-test for heterogeneity and the $I^2$ statistic are reported. We complement this with an assessment of heterogeneity of effect sizes graphically using forest plots. We identified outliers using studentized residuals and identified overly influential studies using Cook’s distance. Where outliers were indicated, we report the resulting effect sizes when they are left out of the analysis. As an additional sensitivity test, we ran a full leave-one-out analysis for all models, and we report these results when and where they are useful. Whenever feasible, we conducted moderator analyses using meta-regression to investigate sources of heterogeneity. (All but two moderators were chosen a priori. Baseline coverage and vaccine hesitancy were added after feedback from an initial peer review from the Campbell Collaboration (copublisher of this work).) The analysis was carried out using R (V.4.0.4) and the metafor package (V.2.4.0). All analyses used a random effects model because we did not reasonably expect the included studies to be functionally identical and the goal was to generalise to the larger population.

Qualitative analysis followed a mix of inductive and deductive coding approaches to identify themes related to barriers and facilitators, reasons for intervention success or failure, and uptake and fidelity challenges. An initial set of themes was developed based on familiarity with the literature. However, as new topics were identified, new themes were added. Themes were also disaggregated if it became clear they were too broad. Research consultants conducted coding with oversight from CL and AB.

**Patient and public involvement in research**

There was no patient or public involvement in this research.

### Results

Our search identified over 43000 records, which were reduced to 29481 unique abstracts after deduplication (figure 1). After title and abstract screening, we considered 1285 studies for full-text screening and could not locate an additional 44, published mostly before 2000. We excluded articles at full-text for not satisfying the inclusion criteria by country (129), study type (evaluation study) (304), evaluation method (213), outcome (255) and community engagement type (172). We ultimately identified 61 impact evaluations (table 2) that assessed the effects of community engagement interventions on outcomes related to routine child immunisation in LMICs. We identified one publication in Spanish, with all others in English. Five studies did not include sufficient data to calculate an effect size, thus, 56 studies were included in the meta-analysis. Inter-rater reliabilities were calculated on a sample of studies, and ranged from 28% (mean effect of the intervention to 100% (eg, country, publication year and study design). All studies were reconciled prior to analyses.

**Risk of bias**

Of the 31 included studies with experimental designs, only two had a low risk of bias, six had some concerns and 23 had a high risk of bias. Of the 30 included quasi-experimental studies, only 2 were assessed as low risk of bias, 1 as some concerns and 27 as high risk of bias. Biases arising from outcome measurement and deviations from intended interventions were the most common across both the study designs.

Although only five qualitative studies scored strong on all key elements, most studies received strong scores on most key elements and had quality appraisal scores greater than 20, indicating low risk of bias. The most common elements found to be missing were sample characteristics and analytical methods. The quality of the cost evidence in the 22 studies that included such evidence was mixed (further information and visualisations for the risk of bias appraisals can be found in online supplemental appendix 8).
Community engagement interventions

Full immunisation

A total of $k = 28$ studies examined the effect of community engagement interventions on full childhood immunisation and found $\hat{\mu} = 0.14$ (95% CI 0.06 to 0.23), $z = 3.28$, $p = 0.01$, indicating a small but significant benefit for the treated group of 0.14 SD units (figure 2). A 95% credibility/prediction interval for the true outcomes is given by $-0.28$ to $0.57$. Hence, although the average outcome is estimated to be positive, in some studies the true outcome may in fact be negative.

The rank correlation test indicated funnel plot asymmetry ($p = 0.03$) but the regression test did not ($p = 0.57$; see online supplemental appendix 9 for additional information). The true effects appear to be heterogeneous ($I^2 = 94.5\%$, $\tau = 0.21$). Outlier analyses revealed that Banerjee may be a potential outlier, and sensitivity analyses removing Banerjee (2010) reduced the overall average effect ($\hat{\mu} = 0.08$ (95% CI 0.04 to 0.12)), but it was still positive and significant ($z = 4.12$, $p < 0.001$). For all other outcomes, outlier analyses will be presented in online supplemental appendices.) Sensitivity analysis using the leave-one-out approach indicates there are no other studies whose removal results in substantial changes to the average effect or overall heterogeneity.

Sensitivity analyses were also conducted to examine the robustness of the results to the exclusion of low-quality studies. When studies assessed as high risk of bias were removed, the resulting effect was slightly larger and still statistically significant ($\hat{\mu} = 0.18$ (95% CI 0.08 to 0.27)), $k = 4$, $z = 3.67$, $p < 0.001$. We examined several potential sources of heterogeneity, including exposure to the intervention, evaluation period, study design, year, geographical region, data source, whether the intervention was implemented by a government agency (either alone or in combination with another agency), whether new cadres of health workers were established, presence of vaccine hesitancy and baseline vaccine coverage rates. There were no significant moderators in the context of this model (see online supplemental appendix 9).

DPT 3

A total of $k = 22$ studies examined the effect on DPT3 vaccination coverage and found a small but significant benefit to the treated group compared with the untreated group ($\hat{\mu} = 0.10$ (95% CI 0.06 to 0.14), $z = 4.75$, $p < 0.001$; figure 3). A 95% credibility/prediction interval for the true outcomes is given by $-0.06$ to $0.26$. Hence, although the average outcome is estimated to be positive, in some studies the true outcome may in fact be negative.
### Table 2 Describing characteristics of the included studies

<table>
<thead>
<tr>
<th>Author</th>
<th>Country</th>
<th>Summary of intervention</th>
<th>Engagement type</th>
<th>Study design</th>
<th>Duration (in months)</th>
<th>Outcomes</th>
</tr>
</thead>
<tbody>
<tr>
<td>Admassie et al</td>
<td>Ethiopia</td>
<td>Formation of a cadre of community-based health extension workers and using community resources for construction of health posts. The study evaluates intervention effectiveness and cost-effectiveness.</td>
<td>Multiple (EII+EAI)</td>
<td>PSM</td>
<td>48</td>
<td>FIC, BCG, DPT3, OPV1, OPV2, OPV3, measles, morbidity</td>
</tr>
<tr>
<td>Adamu et al</td>
<td>Nigeria</td>
<td>Quality improvement programme where health workers use iterative processes to develop localised and contextually relevant plans to resolve health service delivery and demand bottlenecks. The study evaluates intervention effectiveness.</td>
<td>Multiple (EID+EII)</td>
<td>ITS</td>
<td>1</td>
<td>Dropouts</td>
</tr>
<tr>
<td>Alhassan et al</td>
<td>Ghana</td>
<td>Using a bottom-up approach, the intervention recruited and trained community groups to identify service delivery gaps in healthcare facilities. The study evaluates intervention effectiveness.</td>
<td>Multiple (EII+EAI)</td>
<td>RCT</td>
<td>10</td>
<td>FIC</td>
</tr>
<tr>
<td>Andersson et al</td>
<td>Pakistan</td>
<td>Community dialogues to address barriers to vaccination. The guidelines for the dialogue were created after consultation with the intended beneficiaries. The study evaluates intervention effectiveness and cost-effectiveness.</td>
<td>Multiple (EID+EAI)</td>
<td>RCT</td>
<td>8</td>
<td>Knowledge about immunisation, attitude about immunisation, community norms, readiness to vaccinate, household norms and decision-making measles, DPT3</td>
</tr>
<tr>
<td>Arifeen et al</td>
<td>Bangladesh</td>
<td>Formation of a cadre of village health volunteers and enlisting support of local religious leaders to convey messages about child health. The study evaluates intervention effectiveness.</td>
<td>EAI</td>
<td>RCT</td>
<td>71</td>
<td>Measles, mortality</td>
</tr>
<tr>
<td>Assegaai et al</td>
<td>South Africa</td>
<td>Lay community-based workers were formalised as community health workers and served as a part of the outreach teams. The study evaluates intervention effectiveness.</td>
<td>EAI</td>
<td>DID</td>
<td>36</td>
<td>FIC, measles, morbidity</td>
</tr>
<tr>
<td>Banerjee et al</td>
<td>India</td>
<td>Provision of immunisation services and incentives to caregivers. A trusted community-based organisation was a key stakeholder in design and delivery of the intervention. The study evaluates intervention effectiveness and cost-effectiveness.</td>
<td>Treatment 1: EAI</td>
<td>RCT</td>
<td>18</td>
<td>FIC, BCG, partial immunisation</td>
</tr>
<tr>
<td>Banerjee et al</td>
<td>India</td>
<td>This evaluation tested two different interventions: 1. Incentives to caregivers. The community’s feedback was solicited on the kind of incentive. 2. Community influencers were identified to spread information about immunisation. The study evaluates intervention effectiveness and cost-effectiveness.</td>
<td>EID EAI</td>
<td>RCT</td>
<td>14</td>
<td>Knowledge about immunisation, attitude about immunisation, FIC, DPT1, DPT2, DPT3, measles</td>
</tr>
<tr>
<td>Banwat et al</td>
<td>Nigeria</td>
<td>Female members of the community whose children are fully immunised were nominated in each community to serve as peer educators. The study evaluates intervention effectiveness.</td>
<td>EAI</td>
<td>CBA</td>
<td>–</td>
<td>FIC, knowledge about immunisation, attitude about immunisation, readiness to vaccinate</td>
</tr>
<tr>
<td>Biemba et al</td>
<td>Zambia</td>
<td>A national policy to create a cadre of well-trained and motivated community-based health workers. The study evaluates intervention effectiveness.</td>
<td>EAI</td>
<td>DID</td>
<td>23</td>
<td>FIC, morbidity</td>
</tr>
<tr>
<td>Björkman et al</td>
<td>Uganda</td>
<td>Communities were involved in monitoring the quality of health services and the performance of health service providers. The study evaluates intervention effectiveness and cost-effectiveness.</td>
<td>EII</td>
<td>RCT</td>
<td>0.16</td>
<td>FIC, BCG, OPV0, OPV3, DPT1, DPT3, measles, Partial routine immunisation, mortality</td>
</tr>
<tr>
<td>Bolam et al</td>
<td>Nepal</td>
<td>Training for community health workers and midwives which was developed in collaboration with health workers and experts. The study evaluates intervention effectiveness.</td>
<td>EID</td>
<td>RCT</td>
<td>3</td>
<td>FIC</td>
</tr>
<tr>
<td>Borkum et al</td>
<td>India</td>
<td>Performance-based incentives to frontline workers. The nature of incentives was decided on in consultation with the frontline workers. The study evaluates intervention effectiveness and cost-effectiveness.</td>
<td>EID</td>
<td>RCT</td>
<td>12</td>
<td>Health card availability, CHW capacity, FIC, BCG, DPT1, DPT2, DPT3, OPV1, OPV2, OPV3, measles, partial immunisation, timeliness</td>
</tr>
<tr>
<td>Calderón-Ortiz and Mejia-Mejía</td>
<td>Mexico</td>
<td>Creation of a community-based cadre of volunteers to register and track children in the community for immunisation. The study evaluates intervention effectiveness.</td>
<td>EAI</td>
<td>CBA</td>
<td>4</td>
<td>FIC, BCG, DPT3, OPV3, measles</td>
</tr>
<tr>
<td>Camell et al</td>
<td>Ethiopia</td>
<td>Formation of a cadre of community health workers to mobilise the community and encourage uptake of health services. The study evaluates intervention effectiveness and cost-effectiveness.</td>
<td>EAI</td>
<td>DID</td>
<td>60</td>
<td>DPT3, measles</td>
</tr>
</tbody>
</table>
## Table 2

Continued

<table>
<thead>
<tr>
<th>Author</th>
<th>Country</th>
<th>Summary of intervention</th>
<th>Engagement type</th>
<th>Study design</th>
<th>Duration (in months)</th>
<th>Outcomes</th>
</tr>
</thead>
<tbody>
<tr>
<td>Costa-Font et al 2017</td>
<td>India</td>
<td>Establishment of the village health and sanitation committees to monitor health service provision at the community level. The study evaluates intervention effectiveness.</td>
<td>EAI</td>
<td>IV</td>
<td>–</td>
<td>BCG, DPT1, OPV0</td>
</tr>
<tr>
<td>Demilew et al 2020</td>
<td>Ethiopia</td>
<td>A poster/stamp system that reminded health workers of the child's immunisation status and simultaneously encouraged caregivers to immunise their children. The intervention was designed in consultation with health workers. The study evaluates intervention effectiveness and cost-effectiveness.</td>
<td>Multiple (EID+EAI)</td>
<td>RCT</td>
<td>17</td>
<td>FIC, BCG, DPT1, DPT2, DPT3, partial immunisation</td>
</tr>
<tr>
<td>Dipeolu 2017</td>
<td>Nigeria</td>
<td>Text message reminders to mothers regarding immunisation schedule. The messages were field tested with mothers to get the content right. The study evaluates intervention effectiveness.</td>
<td>EID</td>
<td>DID</td>
<td>9</td>
<td>Knowledge about immunisation, attitude about immunisation, timeliness</td>
</tr>
<tr>
<td>Domek et al 2019</td>
<td>Guatemala</td>
<td>SMS text messages to caregivers. A prior feasibility and acceptability study was conducted for the intervention. The study evaluates intervention effectiveness.</td>
<td>EID</td>
<td>RCT</td>
<td>2</td>
<td>Timeliness</td>
</tr>
<tr>
<td>Engineer et al 2016</td>
<td>Afghanistan</td>
<td>Pay-for-performance bonuses paid quarterly to health workers. The bonus amount was revised after receiving health worker feedback. The study evaluates intervention effectiveness.</td>
<td>EID</td>
<td>RCT</td>
<td>24</td>
<td>Experience &amp; satisfaction with health services, Formal HW motivation, capacity &amp; performance, DPT3</td>
</tr>
<tr>
<td>Findley et al 2015</td>
<td>Nigeria</td>
<td>Formation of a cadre of community volunteers to facilitate group discussions on health and track/register women/children for health services. The study evaluates intervention effectiveness and cost-effectiveness.</td>
<td>EAI</td>
<td>DID</td>
<td>24</td>
<td>FIC</td>
</tr>
<tr>
<td>Gibson et al 2017</td>
<td>Kenya</td>
<td>SMS reminders and monetary incentives to caregivers. A feasibility study was conducted in 2013 for this intervention. The study evaluates intervention effectiveness.</td>
<td>EID</td>
<td>RCT</td>
<td>12</td>
<td>Community norms, FIC, BCG, DPT1, DPT2, DPT3, OPV0, OPV1, OPV2, measles, timeliness</td>
</tr>
<tr>
<td>Goel et al 2012</td>
<td>India</td>
<td>A multi-component campaign which involved women groups in awareness generation to improve health service uptake. The study evaluates intervention effectiveness.</td>
<td>EAI</td>
<td>DID</td>
<td>48</td>
<td>FIC</td>
</tr>
<tr>
<td>Gurley et al 2020</td>
<td>India</td>
<td>Community members were trained to design and produce culturally appropriate, ‘hyperlocal’ videos to promote health seeking behaviours. The study evaluates intervention effectiveness.</td>
<td>Multiple (EID+EAI)</td>
<td>RCT</td>
<td>11</td>
<td>Knowledge about immunisation, attitude about immunisation, FIC, DPT3, partial immunisation, timeliness, dropouts</td>
</tr>
<tr>
<td>Herrera-Almanza and Rosales-Rueda 2018</td>
<td>Madagascar</td>
<td>Community-based primary healthcare services intervention that included the deployment of volunteer community health workers in remote areas. The study evaluates intervention effectiveness.</td>
<td>Multiple (EII+EAI)</td>
<td>DID</td>
<td>26</td>
<td>Health card availability, OPV3, DPT3, measles, mortality, partial immunisation</td>
</tr>
<tr>
<td>Igarashi et al 2010</td>
<td>Zambia</td>
<td>The GMP+sessions were conducted by medical personnel from Public Health Centres. During these sessions, community volunteers provided some operational and managerial support to ensure the effective implementation of the sessions. The study evaluates intervention effectiveness.</td>
<td>EII</td>
<td>CBA</td>
<td>43</td>
<td>FIC, timeliness, attitude about immunisation, community norms</td>
</tr>
<tr>
<td>Janissens 2011</td>
<td>India</td>
<td>Dissemination of health promoting messages to women in the community who are encouraged to further spread the awareness. The study evaluates intervention effectiveness.</td>
<td>EAI</td>
<td>IV</td>
<td>56.4</td>
<td>DPT3, measles</td>
</tr>
<tr>
<td>Johri et al 2020</td>
<td>India</td>
<td>Interventions, designed through formative research, to increase caregiver knowledge and adherence to childhood immunisation. The study evaluates intervention effectiveness.</td>
<td>EID</td>
<td>RCT</td>
<td>3</td>
<td>Knowledge about immunisation, Awareness of place, time, schedule for vacc., attitude about immunisation</td>
</tr>
<tr>
<td>Lee 2015</td>
<td>Zambia</td>
<td>Creation of a new cadre of frontline workers from the community, called community health assistants, to provide primary healthcare services. The study evaluates intervention effectiveness.</td>
<td>EAU</td>
<td>RCT</td>
<td>3</td>
<td>Formal HW motivation, capacity &amp; performance, BCG, OPV3, measles, timeliness, CHW capacity, morbidity</td>
</tr>
<tr>
<td>Mayumana et al 2017</td>
<td>Tanzania</td>
<td>Payment-for-performance scheme for health facilities. Health workers and health facility governing committees decided the allocation of funds. The study evaluates intervention effectiveness.</td>
<td>EII</td>
<td>DID</td>
<td>30</td>
<td>Stockouts</td>
</tr>
</tbody>
</table>

**Continued**
<table>
<thead>
<tr>
<th>Author</th>
<th>Country</th>
<th>Summary of intervention</th>
<th>Engagement type</th>
<th>Study design</th>
<th>Duration (in months)</th>
<th>Outcomes</th>
</tr>
</thead>
<tbody>
<tr>
<td>Memon et al</td>
<td>Pakistan</td>
<td>Formation of community health committees to promote perinatal and new-born care. Formative research informed the intervention design. The study evaluates intervention effectiveness.</td>
<td>Multiple</td>
<td>DID</td>
<td>16</td>
<td>FIC</td>
</tr>
<tr>
<td>Modi et al</td>
<td>India</td>
<td>The mHealth intervention package consisting of mobile phone-based job aids for community health workers. The intervention was piloted in 2015. The study evaluates intervention effectiveness and cost-effectiveness.</td>
<td>EID</td>
<td>RCT</td>
<td>12</td>
<td>DPT3, morbidity</td>
</tr>
<tr>
<td>Mohanan et al</td>
<td>India</td>
<td>Social accountability interventions to promote community-based collective action to improve delivery of health and nutrition services to children. The study evaluates intervention effectiveness and cost-effectiveness.</td>
<td>EAI</td>
<td>RCT</td>
<td>12</td>
<td>Experience &amp; satisfaction with health services, attitudes about health providers, formal health worker supply, FIC, BCG, DPT3, OPV1, OPV3, IPV, measles, morbidity, mortality</td>
</tr>
<tr>
<td>More et al</td>
<td>India</td>
<td>Urban slum-dweller women’s groups used community dialogues to address barriers to improving perinatal health. The study evaluates intervention effectiveness.</td>
<td>EAI</td>
<td>RCT</td>
<td>36</td>
<td>Mortality</td>
</tr>
<tr>
<td>More et al</td>
<td>India</td>
<td>The intervention comprised multiple activities like home visits to caregivers, groups meetings, community events and other supportive services. The study evaluates intervention effectiveness and cost-effectiveness.</td>
<td>EAI</td>
<td>RCT</td>
<td>24</td>
<td>Health card availability, attitudes about health providers, FIC, BCG, measles, partial immunisation</td>
</tr>
<tr>
<td>Morris et al</td>
<td>Honduras</td>
<td>Monetary vouchers to women in the communities and setting up of community-based committees to oversee health service quality and access. The study evaluates intervention effectiveness and cost-effectiveness.</td>
<td>EAI</td>
<td>RCT</td>
<td>24</td>
<td>DPT1, measles</td>
</tr>
<tr>
<td>Murthy et al</td>
<td>India</td>
<td>Voice call reminders to pregnant women and caregivers. The message content was tested for appropriateness through community focused groups. The study evaluates intervention effectiveness.</td>
<td>EID</td>
<td>RCT</td>
<td>21</td>
<td>Knowledge about immunisation, FIC</td>
</tr>
<tr>
<td>Nagar et al</td>
<td>India</td>
<td>A digital pendant-based health record of the child and a voice call reminder system. A formative study was conducted in 2016 and communities were consulted on the design of the pendant. The study evaluates intervention effectiveness.</td>
<td>EID</td>
<td>RCT</td>
<td>3</td>
<td>Timeliness</td>
</tr>
<tr>
<td>Nagar et al</td>
<td>India</td>
<td>A digital pendant-based health record of the child. Health providers used a mobile application to scan the pendant to update the child’s medical history. Prior formative research informed the intervention design. The study evaluates intervention effectiveness and cost-effectiveness.</td>
<td>EID</td>
<td>RCT</td>
<td>20</td>
<td>FIC, DPT1, DPT2, DPT3</td>
</tr>
<tr>
<td>Nzioki et al</td>
<td>Kenya</td>
<td>Formation of a cadre of community health workers. The study evaluates intervention effectiveness.</td>
<td>EAI</td>
<td>CBA</td>
<td>0</td>
<td>FIC</td>
</tr>
<tr>
<td>Oche et al</td>
<td>Nigeria</td>
<td>Group meetings with caregivers and dialogues with community leaders to improve uptake of routine immunisation services. The study evaluates intervention effectiveness.</td>
<td>EAI</td>
<td>DID</td>
<td>9</td>
<td>Knowledge about immunisation, DPT1, DPT3, dropouts</td>
</tr>
<tr>
<td>Okeke et al</td>
<td>Nigeria</td>
<td>A national scheme to create, train and deploy a cadre of midwives to serve underserved rural and remote populations in Nigeria. The study evaluates intervention effectiveness.</td>
<td>Multiple</td>
<td>DID</td>
<td>40</td>
<td>BCG, DPT3, OPV3, measles, mortality</td>
</tr>
<tr>
<td>Okoli et al</td>
<td>Nigeria</td>
<td>A conditional cash transfer programme to encourage uptake of health services. Community groups were consulted while deciding the cash amount. The study evaluates intervention effectiveness.</td>
<td>EID</td>
<td>ITS</td>
<td>seven to 18</td>
<td>OPV0</td>
</tr>
<tr>
<td>Olayo et al</td>
<td>Kenya</td>
<td>Formation of a cadre of community health workers who then facilitated dialogue at the community level and supported other community-based workers. The study evaluates intervention effectiveness.</td>
<td>Multiple</td>
<td>DID</td>
<td>24</td>
<td>Health card availability, DPT1, DPT3, measles</td>
</tr>
<tr>
<td>Olken et al</td>
<td>Indonesia</td>
<td>Block grants for maternal and child health that incorporated relative performance incentives were implemented in villages through creation of village-level health committees. The study evaluates intervention effectiveness and cost-effectiveness.</td>
<td>Multiple</td>
<td>RCT</td>
<td>18 to 30</td>
<td>FIC, morbidity, mortality</td>
</tr>
</tbody>
</table>

Continued
<table>
<thead>
<tr>
<th>Author</th>
<th>Country</th>
<th>Summary of intervention</th>
<th>Engagement type</th>
<th>Study design</th>
<th>Duration (in months)</th>
<th>Outcomes</th>
</tr>
</thead>
<tbody>
<tr>
<td>Oyo-Ita et al.</td>
<td>Nigeria</td>
<td>A multi-component intervention involving traditional and religious leaders for engaging</td>
<td>Multiple (EII+EAI)</td>
<td>RCT</td>
<td>18</td>
<td>FIC, partial immunisation, Timeliness</td>
</tr>
<tr>
<td>Rahman et al.</td>
<td>India</td>
<td>Trained facilitators from local NGOs interacted with the communities to enable them to</td>
<td>EAI</td>
<td>RCT</td>
<td>13</td>
<td>Knowledge about immunisation, attitude about immunisation, attitudes</td>
</tr>
<tr>
<td>Rahman et al.</td>
<td>Pakistan</td>
<td>Mental health support programme with counselling sessions for pregnant and post-partum</td>
<td>EID</td>
<td>RCT</td>
<td>11</td>
<td>FIC, morbidity</td>
</tr>
<tr>
<td>Rao 2014</td>
<td>India</td>
<td>Creation of a cadre of community health workers to improve basic health outcomes through</td>
<td>EAI</td>
<td>DID</td>
<td>60</td>
<td>BCG, DPT3, OPV3, measles, FIC, partial immunisation, supply of CHWs</td>
</tr>
<tr>
<td>Roy et al. 2008</td>
<td>Bangladesh</td>
<td>New cadre of community health workers delivered essential maternal, neonatal and child</td>
<td>EAI</td>
<td>DID</td>
<td>11</td>
<td>FIC, BCG, DPT1, DPT2, DPT3, OPV0, OPV1, OPV2, OPV3, measles, partial</td>
</tr>
<tr>
<td>Saggurti et al.</td>
<td>India</td>
<td>Formation of health-focused self-help groups with women of reproductive age coming from</td>
<td>EAI</td>
<td>DID</td>
<td>2</td>
<td>Timeliness</td>
</tr>
<tr>
<td>Sankar 2013</td>
<td>India</td>
<td>Formation of committees with representatives of the community, local government and</td>
<td>Multiple (EII+EAI)</td>
<td>DID</td>
<td>30</td>
<td>FIC, BCG, DPT1, DPT2, DPT3, OPV0, OPV1, OPV2, OPV3, measles, partial</td>
</tr>
<tr>
<td>Seth et al. 2018</td>
<td>India</td>
<td>The study evaluated two different interventions: role of compliance-linked incentives vs</td>
<td>EID</td>
<td>RCT</td>
<td>9.7</td>
<td>Partial immunisation, timeliness, attitude about immunisation, attitudes</td>
</tr>
<tr>
<td>Shukla 2018</td>
<td>Afghanistan</td>
<td>Community representatives along with health officials identify the health needs of the</td>
<td>EII</td>
<td>DID</td>
<td>6</td>
<td>Supply of CHWs, DPT3</td>
</tr>
<tr>
<td>Siddique et al.</td>
<td>Pakistan</td>
<td>Visual reminders to caregivers in the form of wearable bracelets for the child. The</td>
<td>EID</td>
<td>RCT</td>
<td>12</td>
<td>DPT3, measles</td>
</tr>
<tr>
<td>Tandon et al.</td>
<td>India</td>
<td>Enlisting community-based volunteers to motivate and encourage family members to use</td>
<td>EAI</td>
<td>CBA</td>
<td>120</td>
<td>FIC, BCG, DPT2, DPT3, OPV2</td>
</tr>
<tr>
<td>USAID 2008</td>
<td>Ethiopia</td>
<td>Creation of a cadre of community health promoters to carry out behaviour change</td>
<td>EAI</td>
<td>CBA</td>
<td>48</td>
<td>FIC, BCG, DPT1, DPT3, OPV3, morbidity, health card availability, measles,</td>
</tr>
</tbody>
</table>

Continued
The true outcomes appear to be heterogeneous ($I^2=76.8\%$, $\tau=0.08$). When low-quality studies were removed, the average effect increased slightly ($\hat{\mu}=0.11$ (95%CI 0.05 to 0.17), $k=4$), and was still statistically significant ($z=3.70$, $p<0.001$; see online supplemental appendix 9). Publication year was a significant source of heterogeneity; each additional year reduced the size of the effect by 0.014 SD units (see online supplemental appendix 9).

### Measles

A total of $k=20$ studies examined the effect on measles vaccination coverage and found a very small but significant benefit for the treated group compared with the untreated group ($\hat{\mu}=0.07$ (95%CI 0.03 to 0.11), $z=3.22$, $p<0.01$; see figure 4). A 95% credibility/prediction interval for the true outcomes is given by $-0.08$ to $0.22$. Hence, although the average outcome

### Table 2

<table>
<thead>
<tr>
<th>Author et al.</th>
<th>Country</th>
<th>Summary of intervention</th>
<th>Engagement type</th>
<th>Study design</th>
<th>Duration (in months)</th>
<th>Outcomes</th>
</tr>
</thead>
<tbody>
<tr>
<td>Webster et al. 2019</td>
<td>Uganda</td>
<td>Community-based outreach and follow-up with caregivers to improve immunisation uptake and reduce defaulters. The study evaluates intervention effectiveness.</td>
<td>EII</td>
<td>RCT</td>
<td>12</td>
<td>Health card availability, BCG, DPT1, DPT2, DPT3, OPV0, OPV1, OPV2, OPV3, IPV, measles, partial immunisation, timeliness, attitude about immunisation, dropouts, morbidity, mortality</td>
</tr>
<tr>
<td>Younes et al. 2014</td>
<td>Bangladesh</td>
<td>The intervention involved 162 women’s groups who used participatory approaches to discuss maternal and neonatal health issues. The study evaluates intervention effectiveness.</td>
<td>EAI</td>
<td>DID</td>
<td>20</td>
<td>FIC, morbidity</td>
</tr>
</tbody>
</table>

CBA, Controlled before-after; CHWs, community health workers; DID, Difference-in-difference; EAI, Engagement as intervention; EID, Engagement in Design; EII, Engagement in implementation autonomy; FIC, Full immunisation coverage; ITS, Interrupted time series; NGOs, Non-governmental organisations; RCT, randomised controlled trial.
is estimated to be positive, in some studies the true outcome may in fact be negative.

When low-quality studies were removed, the average effect increased ($\hat{\mu} = 0.09$, $k = 6$, (95% CI 0.03 to 0.15) and was still statistically significant $z=2.98$, $p=0.003$). The true outcomes appear to be heterogeneous ($I^2=73.6\%$, $\tau = 0.07$). None of the moderators were significant sources of heterogeneity (see online supplemental appendix 9).

Vaccination timeliness

We found a small but significant effect on all three timeliness outcomes: full immunisation timeliness ($\hat{\mu} = 0.15$ (95% CI 0.07 to 0.24, $z = 3.41$), $p<0.001$, 95% prediction interval 0.04 to 0.27; DPT3 timeliness ($\hat{\mu} = 0.09$ (95% CI 0.03 to 0.14), $z = 3.00$, $p<0.01$, 95% prediction interval 0.03 to 0.14) and measles timeliness ($\hat{\mu} = 0.23$ (95% CI 0.14 to 0.32, $z=5.06$, $p<0.001$, 95% prediction interval 0.14 to 0.32. For all timeliness outcomes, tests of heterogeneity were not significant ($p>0.05$). For full immunisation and measles timeliness outcomes, the sensitivity analysis could not be conducted due to an inadequate number of studies. For DPT3, the average effect increased but became non-significant when low quality studies were removed (see online supplemental appendix 9).

Subgroups of community engagement interventions

Studies that used engagement as the intervention had a significant positive effect on full childhood immunisation, DPT3 vaccination and measles vaccination but evidence was insufficient to synthesise measures of vaccination timeliness (table 3).

When studies used community engagement in the design, there was a significant positive effect on full childhood immunisation and measles vaccination but not on DPT3 vaccination. In addition, there was a positive significant effect on timeliness of full childhood immunisation and DPT3 vaccination. No studies using engagement in the design reported on timeliness of measles vaccinations.

For engagement in implementation autonomy, the analysis is based on a limited number of studies and we found no significant effect on either coverage or timeliness outcomes. There were no studies reporting on the timeliness of measles vaccination.

Finally, some studies combined multiple engagement types in their interventions. These interventions had a significant effect on DPT3 vaccination but not on measles vaccination or full childhood vaccination. Evidence was insufficient to synthesise measures of vaccination timeliness.
Secondary outcomes analyses can be found in online supplemental appendices 10–13.

**Qualitative findings**

Programme design characteristics were associated with intervention success or failure across all engagement types. Certain aspects of community engagement itself, such as conducting stakeholder consultations, holding community dialogues or involving community leaders were associated with better immunisation outcomes. Non-engagement intervention design features also affected intervention success. These design features include incentives given to caregivers and leadership and supportive supervision, which improved overall health service delivery and health worker performance. Among the studies that attributed intervention failure to programme design, inadequate duration, frequency or exposure to the intervention were the most notable themes.

The importance of accounting for contextual barriers to or facilitators of immunisation emerged consistently. Limited availability of services, especially insufficient staff and vaccine supply, were dominant barriers to immunisation, affecting outcomes in the early portion of the causal chain. Other common barriers to immunisation included practical barriers faced by caregivers such as costs, largely indirect and logistics (wait time and language barriers) or distance. There was more variation in barriers related to social norms, fear and an understanding of the importance of immunisation by type of engagement. Poor quality of services, including uninviting attitudes of health workers, posed a barrier to immunisation in communities that received engagement as the intervention or were engaged in the design of the intervention.

However, we also found that certain contextual factors could become facilitators of immunisation outcomes, provided an intervention has adequately situated itself to leverage them. Across all engagement types, studies associated caregivers’ awareness and perception of the benefits of vaccination with improved immunisation outcomes. Similarly, availability of health infrastructure and good quality of services were also associated with improved intervention uptake and its impact.

Implementation failures, such as low fidelity, were a common reason for intervention failure. Across all engagement types, interventions did not properly account for existing implementation constraints and practicalities on the ground and were forced to change their implementation plans. Many of these issues were encountered due to uncontrollable factors or invalid theory of change assumptions. For instance, programme design may not have accounted for the unavailability of intended participants due to competing priorities, thereby potentially invalidating the assumption of beneficiary exposure in the causal pathway. Administrative challenges were cited consistently, though their nature varied across engagement types, ranging from technical limitations (such as limited cellphone service) to political constraints to insufficient staffing levels.

---

**Figure 4** Forest plot showing the observed outcomes and the estimate of the random effects model for the impact of community engagement interventions on measles vaccination. Note that # participants is specific to each effect and thus may not reflect the sample size for the full study.
Results were broadly consistent when only qualitative studies with a quality appraisal score greater than 20 were considered. Among the 14 studies for which we could calculate cost-effectiveness, we found that community engagement interventions had a small, but significant, positive effect on all primary immunisation outcomes related to coverage and their timeliness. We also found that certain features of interventions may contribute to their success. These include (A) appropriate intervention design, including building in community engagement features; (B) addressing the cost of vaccines and practicalities on the ground. We found no ubiquitous patterns of heterogeneity among the primary outcomes.

The median intervention cost per treated child per vaccine dose (excluding the cost of vaccines) to increase absolute immunisation coverage by 1% was US$3.68. Among the four types of community engagement interventions, we found that engagement as the intervention (embedded community engagement), which involves creating community buy-in or development of new community-based structures or cadres, had consistent positive effects on more primary vaccination outcomes than the others. We also found engaging the community in the design of the intervention had a positive effect on most primary outcomes related to coverage. We found no ubiquitous patterns of heterogeneity among the primary outcomes.

While zero dose children were not the specific focus of the review, we can offer some insights based on our analyses of both DPT and BCG outcomes, which reflect access to initial dosing. Community engagement interventions did not show an effect on DPT, but the evidence base was of low quality, with six of eight studies assessed as having a high risk of bias. There was a small but significant effect of community engagement interventions on BCG, but here again, the evidence base was of lower quality, with 9 of 12 studies assessed as having a high risk of bias. In both cases, the evidence base was smaller in size than for the primary outcomes. As we find positive effects of community engagement interventions on children returning for DPT and measles doses, it may be that barriers to vaccination, and specifically the need for community engagement, are more relevant for children who have not received their initial doses.

**Table 3** Overview of the primary outcomes by engagement type

<table>
<thead>
<tr>
<th>Engagement as the intervention</th>
<th>Engagement in the design</th>
<th>Engagement in the implementation autonomy</th>
<th>Multiple engagement types</th>
</tr>
</thead>
<tbody>
<tr>
<td>Random effects summary estimate for all interventions</td>
<td>Engagement as the intervention</td>
<td>Engagement in the design</td>
<td>Engagement in the implementation autonomy</td>
</tr>
<tr>
<td>$g$, (95% CI), $t^2$, (k)</td>
<td>$g$, (95% CI), $t^2$, (k)</td>
<td>$g$, (95% CI), $t^2$, (k)</td>
<td>$g$, (95% CI), $t^2$, (k)</td>
</tr>
<tr>
<td>Full immunisation</td>
<td>0.14*, (0.06 to 0.23), 94.46, (28)</td>
<td>0.08***, (0.03 to 0.13), 70.00, (12)</td>
<td>0.01*, (0.02 to 0.19), 23.83, (5)</td>
</tr>
<tr>
<td>Measles</td>
<td>0.07**, (0.03 to 0.11), 73.64, (20)</td>
<td>0.10**, (0.05 to 0.15), 60.29, (10)</td>
<td>0.11*, (0.02 to 0.21), 0.00, (2)</td>
</tr>
<tr>
<td>DPT3</td>
<td>0.10**, (0.06 to 0.14), 76.78, (22)</td>
<td>0.09**, (0.03 to 0.15), 73.17, (6)</td>
<td>0.04, (−0.01 to 0.08), 0.00, (6)</td>
</tr>
<tr>
<td>Timeliness (DTP3)</td>
<td>0.09**, (0.03 to 0.14), 0.00, (7)</td>
<td>N/A</td>
<td>0.12**, (0.03 to 0.21), 0.00, (4)</td>
</tr>
<tr>
<td>Timeliness (Measles)</td>
<td>0.23**, (0.14 to 0.32), 0.00, (3)</td>
<td>N/A</td>
<td>N/A</td>
</tr>
<tr>
<td>Timeliness (full immunisations)</td>
<td>0.15**, (0.07 to 0.24), 9.65, (5)</td>
<td>N/A</td>
<td>0.15**, (0.004 to 0.29), 7.202, (2)</td>
</tr>
</tbody>
</table>

*p<0.05, **p<0.01, ***p<0.001.
NA, not available.

**DISCUSSION**

**Principal findings**

We found that community engagement interventions had a small, but significant, positive effect on all primary immunisation outcomes related to coverage and their timeliness. We also found that certain features of interventions may contribute to their success. These include (A) appropriate intervention design, including building in community engagement features; (B) addressing the cost of vaccines and practicalities on the ground.

We found no ubiquitous patterns of heterogeneity among the primary outcomes.

While zero dose children were not the specific focus of this review, we can offer some insights based on our analyses of both DPT and BCG outcomes, which reflect access to initial dosing. Community engagement interventions did not show an effect on DPT, but the evidence base was of low quality, with six of eight studies assessed as having a high risk of bias. There was a small but significant effect of community engagement interventions on BCG, but here again, the evidence base was of lower quality, with nine of 12 studies assessed as having a high risk of bias. In both cases, the evidence base was smaller in size than for the primary outcomes. As we find positive effects of community engagement interventions on children returning for DPT and measles doses, it may be that barriers to vaccination, and specifically the need for community engagement, are more relevant for children who have not received their initial doses.

**Table 3** Overview of the primary outcomes by engagement type

<table>
<thead>
<tr>
<th>Random effects summary estimate for all interventions</th>
<th>Engagement as the intervention</th>
<th>Engagement in the design</th>
<th>Engagement in the implementation autonomy</th>
<th>Multiple engagement types</th>
</tr>
</thead>
<tbody>
<tr>
<td>$g$, (95% CI), $t^2$, (k)</td>
<td>$g$, (95% CI), $t^2$, (k)</td>
<td>$g$, (95% CI), $t^2$, (k)</td>
<td>$g$, (95% CI), $t^2$, (k)</td>
<td>$g$, (95% CI), $t^2$, (k)</td>
</tr>
<tr>
<td>Full immunisation</td>
<td>0.14*, (0.06 to 0.23), 94.46, (28)</td>
<td>0.08***, (0.03 to 0.13), 70.00, (12)</td>
<td>0.01*, (0.02 to 0.19), 23.83, (5)</td>
<td>0.23, (−0.001 to 0.47), 73.07, (2)</td>
</tr>
<tr>
<td>Measles</td>
<td>0.07**, (0.03 to 0.11), 73.64, (20)</td>
<td>0.10**, (0.05 to 0.15), 60.29, (10)</td>
<td>0.11*, (0.02 to 0.21), 0.00, (2)</td>
<td>0.03, (−0.09 to 0.15), 54.84, (2)</td>
</tr>
<tr>
<td>DPT3</td>
<td>0.10**, (0.06 to 0.14), 76.78, (22)</td>
<td>0.09**, (0.03 to 0.15), 73.17, (6)</td>
<td>0.04, (−0.01 to 0.08), 0.00, (6)</td>
<td>0.11, (−0.05 to 0.28), 80.12, (3)</td>
</tr>
<tr>
<td>Timeliness (DTP3)</td>
<td>0.09**, (0.03 to 0.14), 0.00, (7)</td>
<td>N/A</td>
<td>0.12**, (0.03 to 0.21), 0.00, (4)</td>
<td>0.04, (−0.06 to 0.13), 0.00, (2)</td>
</tr>
<tr>
<td>Timeliness (Measles)</td>
<td>0.23**, (0.14 to 0.32), 0.00, (3)</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
</tr>
<tr>
<td>Timeliness (full immunisations)</td>
<td>0.15**, (0.07 to 0.24), 9.65, (5)</td>
<td>N/A</td>
<td>0.15**, (0.004 to 0.29), 7.202, (2)</td>
<td>0.38, (−0.28 to 1.03), 0.00, (2)</td>
</tr>
</tbody>
</table>

*p<0.05, **p<0.01, ***p<0.001.
NA, not available.
like availability of health services, for zero dose children are different and unless those are addressed, community engagement itself may not be enough.

**Strengths and weakness**
Our systematic review uses a detailed framework of community engagement interventions to assess their effectiveness for improving outcomes related to routine child immunisation in LMICs. As far as we are aware, ours is the first systematic review to do this. Sensitivity analyses excluding high risk of bias studies showed that the effect was slightly larger and still statistically significant for almost all the primary outcomes for which we had sufficient data. The effects were also uniform across geographies and baseline immunisation rates.

We drew on 61 studies for meta-analysis, comprising 31 RCTs and 30 quasi experimental studies. However, only 56 studies provided sufficient information for calculating effect sizes and thus were included in meta-analytical models. For full immunisation, DPT3 and measles coverage, we could draw on 28, 22 and 20 studies, respectively, for pooled effects. However, for the timeliness of these coverage outcomes we had only 0–7 studies to assess the pooled effects. Thus, while for some outcomes the evidence base for drawing conclusions is adequate, for others it is limited. Among the four kinds of community engagement interventions, there was a relatively large evidence base for those with engagement as the intervention and those with multiple engagement types, while for interventions using engagement in implementation autonomy, the evidence was quite limited.

We identified additional documentation comprising qualitative studies, project reports, formative/ process evaluations and observation studies for 39 of the 61 included impact evaluations. However, the crucial qualitative papers which help us gain a deeper understanding of overall intervention mechanisms of change were found for only 17 of the 61 IEs. Likewise, only 14 of the primary studies included in this review both estimated the intervention cost and reported it with sufficient detail for the review team to calculate the cost-effectiveness of the treatment. Low-quality cost data and the unavailability of underlying cost data contributed to the small number of primary studies included in the cost-effectiveness analysis.

The quantitative evidence was mostly low quality, though the randomised studies were generally of higher quality and less likely to have confounding bias than the quasi-experimental studies. The quality of qualitative studies was generally high. The quality of the cost evidence was mixed. Despite the quality concerns about quantitative evidence, the sensitivity analysis conducted by excluding low-quality studies corroborated the overall findings. Despite a comprehensive search strategy and the inclusion of grey literature, publication bias was detected for the three primary coverage outcomes (full immunisation, DPT3 and measles). While bias correction analyses indicated an identical effect size for full immunisation and DPT3, the effect for measles was reduced when publication bias was corrected for. Timeliness outcomes had an insufficient number of studies to test for publication bias, which limits our ability to interpret heterogeneity.

**Agreement and disagreement with other reviews**
The findings from this review are broadly consistent with Molina et al., which found positive effects of community monitoring interventions on immunisation coverage. Another review by Gilmore and McAuliffe examined the effectiveness of preventive interventions delivered by community health workers for maternal and child health in LMICs on essential newborn care and found some evidence in its support through narrative synthesis, but found the evidence base to be insufficient to draw firm conclusions.

**Limitations**
There are several potential limitations to the current review: (A) there were few analyses that were sufficiently powered to test for publication bias, thus, we cannot rule this out in many cases; (B) many of the moderator analyses were underpowered, meaning that in many cases we were unable to explore heterogeneity. This was particularly true in the context of the subgroup analyses of the four intervention types. In addition, it is likely that there is interdependency among moderator variables, but the current study did not allow for us to disentangle these confounds. Future studies may aim to better assess how moderators may work in tandem to affect the magnitude of change; (C) even in cases where the average effect was significant, forest plots demonstrate that some of the included studies reported a small negative affect, and prediction intervals often included both positive and negative values, which may have important implications when making decisions related to programme design and implementation; (D) we also observed very few studies which focused on subpopulation groups. This is particularly problematic given the focus on LMICs, where equity is important to consider when trying to increase coverage; (E) most of the community engagement interventions were in combination with other intervention components, thus we were not able to establish their unique contribution to changes in outcomes and (F) inclusion of primary studies into this review was based on the description of the community engagement aspects of the intervention. We may have excluded studies that should ideally have been included because of inadequate reporting of intervention components. Finally, funding for this project has concluded, thus, we do not have the resources to update our literature search last conducted in May of 2020.

**Implications for policy and practice**
COVID-19 has impacted routine child immunisation negatively in some countries, and community engagement interventions could be an effective way to counteract this decline. The positive effects of community engagement interventions can be expected across a variety
of settings, although some engagement approaches appear to be more effective than others. Positive design features should be integrated into these interventions, including features such as holding community dialogues or involving community leaders, and non-community engagement features such as local supportive supervision and incentives to healthcare workers or caregivers. Wherever possible, binding contextual barriers to immunisation, such as weak health systems and social norms, should be accounted for in the design of interventions. Existing contextual facilitators for immunisation, such as good existing health systems or high maternal education, could be leveraged for increasing intervention impacts. Important implementation preconditions, such as regular internet service or sufficient staffing, should be assessed and established before the implementation or addressed through the design itself. Close monitoring of intervention implementation along with good understanding of context is important to help make necessary modifications in case of unexpected challenges, such as political instability.

Further research

For better-quality evidence and deeper mechanistic understanding, policy makers and practitioners should consider prioritising funding or commissioning research in the following areas: (A) ways of ameliorating outcome measurement bias due to self-reported immunisation coverage outcomes, as this was a principal source of bias; (B) better reporting of interventions, more rounded analysis of why the interventions worked through mixed-methods evaluations and greater focus on intermediate outcomes for improved understanding of causal mechanisms; (C) collection and reporting of high-quality cost data to enable cost-effectiveness analysis, which is important for decision-making within budget constraints and (D) focus on subgroup analysis, including for zero dose children, for ensuring immunisation services for the most marginalised children. It would also be useful to conduct an update of this review to include evidence produced since our final literature search in May of 2020.

Acknowledgements

This research has been undertaken as a part of 3ie’s immunisation evidence programme, supported by the Bill and Melinda Gates Foundation, Seattle, USA. We would like to thank Sahil Agha, Senior Program Officer, Gates Foundation, Seattle, for his continued support and engagement on the evidence programme and the review. Special thanks to Molly Abruzzese whose guidance helped shape the scope of our evidence programme and aided the conceptualisation of this review. We thank Lisa Menning, Danielle Pedi, Jennifer (Fluder) Siler and Pedja Stocijic for their valuable feedback on the review. We wish to thank John Eyers, Yoav Vardy, Daniela Andia, Shraddha Parsekar, Sejal Luthra, Yue Zhan, Aditi Hombali, Lalitha Vadrevu, Reva Datar, Pankhuri Jha, Beáta Berkovics, Ashton Baafi, Moital Kuperf David Atika, Himani Aggarwal, Agrima Sahore, Mansi Wadhwa, and Harini Narayanan for excellent support and research assistance. We would further like to thank Marie Gaarder, Executive Director, Birte Snilstevit, Director of Synthesis and Reviews Office and Sebastian Martinez, Director of Evaluation Office, 3ie, for their guidance on the review process.

Contributors

MJ, ME and AB conceived the review and wrote the initial protocol. ME and AB did the systematic search. ME, MJ and AB screened and identified studies and MJ and AB made final decisions regarding study inclusion. SS did the statistical analysis. CL and AB did the qualitative analysis. EB synthesised the cost evidence. MJ provided critical inputs on the whole analysis, checked data, coordinated the review and had full access to all materials and results. External consultants supported the authors in search, screening, data extraction and critical appraisal of quantitative and qualitative evidence base. All authors critically reviewed and revised the manuscript and approved the final document for submission. MJ is responsible for the overall content and is the guarantor.

Funding

This research was funded through a grant from the Bill & Melinda Gates Foundation (OPP1115129).

Disclaimer

The funder of the study had no role in study design, data collection, data analysis, data interpretation, or writing of the report.

Competing interests

The International Initiative for Impact Evaluation (3ie).

Through this grant, 3ie provided funding and technical assistance for seven impact evaluations of community engagement interventions for immunisation as a part of its immunisation evidence programme. This technical assistance included, but was not limited to: reviewing study designs, analysis plans and data collection instruments; advising research teams on how to improve study components and address challenges that arise during the course of the evaluation; and supporting grantees in engaging with stakeholders to promote uptake and use of evidence generated by the evaluations. As members of 3ie staff, authors MJ, AB and ME have all had varying levels of involvement in reviewing deliverables for these evaluations and providing research teams with technical assistance. Several procedural safeguards and transparency measures were put in place to mitigate the risk this conflict of interest imposed. First, all candidate studies, including those funded by 3ie, underwent a rigorous multi-step screening process, including review at the title, abstract, and full-text levels. To qualify for inclusion in the SR, a study was judged to meet the inclusion criteria related to study design, outcomes and population by two independent screeners who have reviewed the full text of the study. The 3ie study authors were responsible for assessing whether the studies met the inclusion criteria for community engagement because of the complexity of the framework. However, these authors have no financial interest in this area and have not published any prior reviews on the topic. The remaining study authors have no conflicts of interest to declare.

Patient and public involvement

Patients and/or the public were not involved in the design, or conduct, or reporting, or dissemination plans of this research.

Patient consent for publication

Not applicable.

Ethics approval

Ethical consent from an ethics committee or institutional board was not required as this study does not involve human subjects.

Provenance and peer review

Not commissioned; externally peer reviewed.

Data availability statement

Data are available on reasonable request.

Supplemental material

This content has been supplied by the author(s). It has not been vetted by BMJ Publishing Group Limited (BMJ) and may not have been peer-reviewed. Any opinions or recommendations discussed are solely those of the author(s) and are not endorsed by BMJ. BMJ disclaims all liability and responsibility arising from any reliance placed on the content. Where the content includes any translated material, BMJ does not warrant the accuracy and reliability of the translations (including but not limited to local regulations, clinical guidelines, terminology, drug names and drug dosages), and is not responsible for any error and/or omissions arising from translation and adaptation or otherwise.

Open access

This is an open access article distributed in accordance with the Creative Commons Attribution 4.0 Unported (CC BY 4.0) license, which permits others to copy, redistribute, remix, transform and build upon this work for any purpose, provided the original work is properly cited, a link to the licence is given, and indication of whether changes were made. See: https://creativecommons.org/licenses/by/4.0/.

ORCID iDs

Monica Jain http://orcid.org/0000-0001-5428-377X

Charlotte Lane http://orcid.org/0000-0002-7730-6687

Mark Engelbert http://orcid.org/0000-0002-3665-1257

REFERENCES


Pramanik S, Ghosh A, Goswami A. Impacts of the stimulate, Appreciate, learn, and transfer community engagement approach to increase vaccination coverage in Assam, India. 3ie Grantee final report. New Delhi: International Initiative for Impact Evaluation (3ie); 2020.


Younes L, Houveling TA, Azad K, et al. The effect of participatory women’s groups on infant feeding and child health knowledge,
Use of community engagement interventions to improve child immunisation in low- and middle-income countries: a systematic review and meta-analysis

Supplementary material

Table of Contents

| APPENDIX 1: AMENDMENTS TO INFORMATION PROVIDED IN THE PROTOCOL | 3 |
| APPENDIX 2: OVERVIEW OF COMMUNITY ENGAGEMENT FRAMEWORK | 5 |
| APPENDIX 3: SEARCH STRATEGY AND SEARCH TERMS FOR IMPACT EVALUATIONS | 11 |
| APPENDIX 4: LIST OF EXCLUDED IMPACT EVALUATION STUDIES | 16 |
| APPENDIX 5: QUALITATIVE AND COST-EVIDENCE SEARCH STRATEGY | 28 |
| Qualitative search protocol | 28 |
| Cost evidence search protocol | 30 |
| List of included qualitative studies | 30 |
| List of included other studies | 33 |
| APPENDIX 6: DATA EXTRACTION TOOLS | 37 |
| Quantitative data extraction tool | 37 |
| Qualitative data extraction tool | 42 |
| Cost data extraction tool | 57 |
| APPENDIX 7: RISK OF BIAS ASSESSMENT TOOLS | 65 |
| Quantitative risk of bias tool | 65 |
| Qualitative risk of bias tool | 81 |
| Cost evidence risk of bias tool | 89 |
| APPENDIX 8: RISK OF BIAS IN THE INCLUDED STUDIES | 90 |
| Quantitative risk of bias | 90 |
Qualitative risk of bias

Risk of bias in cost effectiveness estimates

APPENDIX 9: QUANTITATIVE RESULTS – ALL COMMUNITY ENGAGEMENT INTERVENTIONS

APPENDIX 10: QUANTITATIVE RESULTS – ENGAGEMENT AS INTERVENTION

APPENDIX 11: QUANTITATIVE RESULTS - COMMUNITY ENGAGEMENT IN INTERVENTION DESIGN

APPENDIX 12: COMMUNITY ENGAGEMENT IN IMPLEMENTATION AUTONOMY OF INTERVENTIONS

APPENDIX 13: MULTIPLE ENGAGEMENT TYPES

APPENDIX 14: QUALITATIVE SYNTHESIS OF RESULTS

Barriers to immunisation

Facilitators of immunisation

Reasons for intervention success

Reasons for intervention failure

Reasons for heterogenous impacts

Uptake and fidelity challenges

APPENDIX 15: SYNTHESIS OF COST EVIDENCE
Appendix 1: Amendments to information provided in the protocol

1. In the report, we have replaced the word ‘participation’ for ‘engagement’, a change that is also reflected in our review title and objectives. Though the IAP2 framework uses the terms “public participation” and “community engagement” interchangeably, we realise that this may lead to confusion as there may be inter researcher differences in how these terms are perceived and used. Therefore, for maintaining consistency, we have used community engagement throughout the report and avoided any references to community participation. From a linguistic perspective, these words are used interchangeability in our review, as is evident from our protocol (please refer to pp. 1-3). Therefore, this change does not alter the scope of our review in any way. In the protocol we had proposed to use the IAP2 spectrum of community engagement for determination of interventions to be included in the review. However, after pilot testing of the IAP2 framework on a range of community engagement interventions, we found that most interventions are based on a more “utilitarian perspective” put forth in Brunton et al. (2017). Thus, for this review we ultimately settled on the community engagement framework which focuses on process of engagement rather than its intensity. Our approach corresponds to some degree to the “extent of engagement” part of the conceptual framework developed in Brunton et al. (2017) and we also kept the spirit of IAP2 framework by including interventions in which engagement goes beyond one-way communication, i.e. beyond the inform level of IAP2, to include some consultation or dialogue with the community or some decision making by the community members. We found this approach to be relatively easier to apply, less subjective, less prone to classification error and potentially useful to practitioners.

2. In the section on objectives, we have clubbed what were originally review questions 3 and 4 in the protocol (p. 5). These questions related to understanding how implementation features and contextual factors such as barriers and facilitators to immunisation, are associated with relative success or failure of community engagement interventions. When we set out to answer these questions through our qualitative synthesis, we realised that all these factors are interlinked and, in various combinations, may influence the outcomes along the causal chain. Therefore, we combined the two objectives and have presented our analysis and results accordingly. This modification to the review objectives does not alter our study’s scope.

3. We had intended to include the relevant Chinese literature in the review. However, because of political disruptions in Hong Kong, it did not materialize.

4. We have added many more secondary outcomes as per our evidence gap map framework and also their categorization has changed. But data was not available in primary studies for most of them, so they could not be analysed in this review.

5. In the protocol we stated that ‘When multiple papers report different results on an identical outcome, we will contact the authors to enquire about the differences and choose the results that more accurately reflect the impact of the intervention as relevant to our research questions. If contacting the author does not yield a clear decision, we will use results from the paper with the latest publication date.’ However, due to the large scope and limited resources of our review, we instead used results from the most recent paper.

6. In the protocol, we stated that ‘Where multiple outcomes are reported from different specifications, we will select the specification with the lowest risk of bias in attributing impact, for example, the most appropriately specified outcomes equation.’ However, given that risk of bias could not be assessed until after inclusion decisions were made, we instead used the authors preferred specification. If this was not explicitly stated, but the authors reported one specification as the “main results” and other specifications as “robustness checks” we used the “main results” specification. If neither of these applied, we used the specification with the most controls.

7. In the protocol we stated that ‘When multiple treatment arms, we will divide the control group by the number of treatment arms.’ Instead, we chose independent effects, choosing either a) the treatment arm
with the most components, b) the treatment arm that was identified by the authors as the one that would be scaled up, or c) when number of components was equal, the arm that was the most cost effective (e.g. choosing unconditional cash transfers (UCTs) over conditional cash transfers as the administrative costs related to UCTs are significantly lower)
Appendix 2: Overview of community engagement framework

The focus of our review is community engagement interventions, which are increasingly being emphasised in international and national policy frameworks as a means to improve immunisation coverage and reach marginalised communities (UNICEF 2018, WHO 2015, WHO 2017b; WHO 2017a).

The most common approach for categorizing community engagement interventions is probably the International Association for Public Participation (IAP2) framework (iap2.org), which identifies five levels of engagement ranging from inform to empower corresponding to increasing community influence over the decisions. This framework corresponds to the “social justice perspective” in Brunton et al. (2017) in which the community engagement is rooted in concerns about social justice, which requires that the health needs are identified by communities themselves and they mobilise themselves into action to make changes within the community. However, after pilot testing of the IAP2 framework on a range of community engagement interventions, we determined that most interventions are based on a more “utilitarian perspective.” These two perspectives of community engagement are very well captured and articulated in Brunton et al. (2017), who in their systematic review of community engagement narratives in public health point out that:

“Historically, interventions to promote health were driven by professionals, with little or no input from the targeted populations; more recently, community engagement has become central to national strategy and guidance for promoting public health, because, from a ‘utilitarian’ point of view, it is thought that more acceptable and appropriate interventions will result, which may result in improved service use and outcomes. Interventions that are based on a utilitarian perspective seek to involve communities in order to improve the effectiveness of the intervention. The intervention itself may be decided upon before the community is invited for its views; or, while the intervention itself is not designed by community members they may be involved in other ways, such as priority setting, or in its delivery. In utilitarian perspectives, health (and other) services reach out to engage particular communities that they have identified require assistance and the intervention is devised within existing policy, practice, and resource frameworks.”

Due to the difference in how the IAP2 framework approaches community engagement and how interventions are actually implemented, it was difficult to systematically categorize interventions using the IAP2 framework with a high degree of consistency among coders. In addition, even for the interventions rooted in “social justice” perspective, their description in the studies was too limited to identify them and map their intensity of community engagement without introducing a lot of subjectivity and non-systematicness.

To avoid this misclassification as far as possible, we experimented with different frameworks and ultimately settled on the one which focuses on process of engagement rather than its intensity. We found that focusing on when and how the community is engaged is a more practical framework for the kind of community engagement interventions that have been evaluated in real world settings. Our approach corresponds to some degree to the “extent of engagement” part of the conceptual framework developed in Brunton et al. (2017) and we also kept the spirit of IAP2 framework by including interventions in which engagement goes beyond one-way communication, i.e. beyond the inform level of IAP2, to include some consultation or dialogue with the community or some decision making by the community members. We found this approach to be relatively easier to apply, less subjective, less prone to classification error and potentially useful to practitioners. The development of this process-oriented framework took a period of around one month and involved three of our core team members.

We consider three points within an intervention during which engagement can occur: engagement can occur in the design of the intervention, engagement can occur in the implementation of the intervention, or the intervention may be engagement. We break these categories into further sub-groups representing specific ways through which
engagement can occur; for example, the development of new cadres of health workers, pilot studies, and the involvement of the community in governance and decision making.

This framework adds a new dimension to the discussion of types of engagement. As practitioners consider the design of their programs, they can use the evidence provided through this framework to determine when and how best to engage with the community.

**Engagement as the intervention (engagement is embedded):** In these interventions a serious attempt was made to gain community buy-in for activities or new cadres of community-based structures were established, such as village health committees or community health volunteers. While, in a few cases, the development of new cadres can be solely for the purpose of didactic teaching and one-way communication, it generally results in dynamic discussion and two-way communication. As such, we chose to include these interventions. Interventions which are themselves community engagement can motivate communities to take ownership of service delivery and address local problems with local solutions.

**Engagement in the design of interventions:** In these interventions community input or feedback was sought before the implementation of an intervention. Such feedback can take the form of a pilot, needs assessment, formative evaluation, or other outreach effort. This form of engagement must occur before the implementer undertakes an action. These interventions allow the community to influence the form of the ultimate action taken. Depending on the weight that is given to the feedback, these engagement activities can align to any of the IAP2 categories other than inform, which was not considered.

**Engagement in implementation autonomy of interventions:** In many interventions the community is not asked for input in their design, but is utilized in their implementation as health care workers, facilitators, or problem solvers. In the spirit of IAP2 framework to go beyond inform level interventions, we only included those interventions in this category where the community members involved in the implementation of the intervention had some opportunity to affect or influence its implementation. This broadly aligns with the involve, collaborate and empower levels of engagement. Due to the inclusion criteria of some autonomy in implementation, the interventions under this category generally involved an existing community led governance structure which weighed in on implementation decisions or the community providing resources without which the intervention could not be implemented. We excluded interventions in which community members, like community health workers or frontline health workers, were involved in implementation, but no new cadres of health workers were created and they could not influence its implementation. For example, community health workers supplying hygiene kits or doing home visits. In addition, interventions which built capacity of existing cadres of community members or provided supportive supervision were excluded. For example, m-health apps for community workers and training of peer facilitators or community health workers or frontline health workers that only allowed for a one-way transfer of knowledge.
Examples of engagement types

ENGAGEMENT AS THE INTERVENTION (ENGAGEMENT IS EMBEDDED)
Effect of health intervention integration within women’s self-help groups on collectivization and healthy practices around reproductive, maternal, neonatal and child health in rural India
Health-focused self-help groups were created for women of reproductive age in marginalized communities.

Effect of peer education on knowledge, attitude and completeness of childhood routine immunization in a rural community of Plateau State.
A new cadre of peer educators was created. Women were trained to provide women with information about routine childhood immunisation. This does not qualify as engagement in implementation because the intervention was the training of peer educators and not the subsequent actions of the peer educators.

Impacts of engaging communities through traditional and religious leaders on vaccination coverage in Cross River State, Nigeria
Traditional and religious leaders were trained to utilize their leadership role to support immunisation. After the training, these leaders then presented data at ward development committee meetings and engaged with the community to encourage immunisation.

ENGAGEMENT IN THE DESIGN OF AN INTERVENTION
Cognitive behaviour therapy-based intervention by community health workers for mothers with depression and their infants in rural Pakistan: a cluster-randomised controlled trial
Many pilots directly seek the feedback of community members. In this study, health workers and depressed mothers were asked about the relevance and usefulness of the intervention before scaling.

Mobile Phone Incentives for Childhood Immunizations in Rural India
Caregivers were given mobile phone credit incentives for completing iminizations. The amount of the incentive was decided upon through conversations that involved community members.

ENGAGEMENT IN THE IMPLEMENTATION AUTONOMY OF AN INTERVENTION
The impact of an immunization programme administered through the Growth Monitoring Programme Plus as an alternative way of implementing Integrated Management of Childhood Illnesses in urban-slum areas of Lusaka, Zambia
The GMP+ sessions were conducted by medical personnel from Public Health Centers. During these session, community volunteers provided some operational and managerial support to ensure the effective implementation of the sessions.

Effects of payment for performance on accountability mechanisms: Evidence from Pwani, Tanzania
The intervention is a performance-based financing mechanism. The community health committee was involved in decisions about how to spend the funds gained through the pay for performance mechanism.

Interventions are defined based on the actions of the external actors, ie implementers, not the community itself. If an intervention spurred the community to take action, the categorization of the intervention is based on the intervention which spurred the action and not the actions that the community took as a result of the intervention. This is because another community might take different actions as a result of the same intervention. The distinction is especially important for interventions to empower communities to improve their own systems. In these cases, the intervention is the activity that empowered the community; this intervention is engagement. For example, interventions which develop village committees to identify local challenges and solutions are themselves engagement interventions. This is because the intervention is the development of the committee and not the actions of the committee. A similar committee in a different area may choose to take different actions, even if the implementing agency does the same development process. Although the community may have engaged in the design of the committee, they likely were not engaged in the design of development process which brought about the committee. For interventions to be qualified as engagement in the design or implementation, the intervention must be an action by an implementer which the community influenced. For example, community members could be consulted in the design of materials used in
outreach activities (Murthy 2019, Nagar 2020). In the example of the development of a new committee, the implementer influenced the action of the community, the community did not influence the action of the implementer.

We defined “communities” in reference to the lowest level of the health service delivery system (or whatever level provides routine immunisation services in the local context). A community is a group of people who are served by a particular primary health facility. Thus, communities encompass a wide range of stakeholders, including caregivers, health service providers, and influential community members such as religious or other traditional leaders. Therefore, our review included any intervention that was directed towards any of these types of community members. Interventions that targeted higher levels of the health system, such as state-level officials, were excluded.

How the interventions might work

A 2015 3ie scoping paper (Sabarwal et al. 2015) systematically mapped the literature on immunisation interventions involving community engagement. Several programme managers and policy experts provided insight regarding why community engagement could be the key to improving immunisation outcomes for children in areas where the coverage has stagnated or declined or that are hard to reach. The findings from the scoping study indicate that working with or engaging communities could help develop a better understanding of the context, target population, problems and barriers, and lead to identification of contextually relevant solutions and desired outcomes, and mobilising community support for them. Because individuals usually function under the influence of social norms, efforts to change these norms can be effective in changing behaviour (Bicchieri and Xiao 2009; Reynolds Subasić and Tindall 2015). People respond to their peers and community and, while activities such as information and education campaigns might have some influence, individuals might feel bound by collective decisions, preventing sustained change (Riedy et al. 2012). The role of peers and of social norms in shaping attitudes towards vaccination is particularly important given that vaccine hesitancy has been documented in countries of all income levels (although it takes different forms in different countries; Dubé, Gagnon, Nickels, Jeram, and Schuster 2014). Community engagement may be effective at overcoming these barriers to immunisation. In both high-income countries (HICs; O'Mara-Eves et al. 2013) and LMICs (De Buck et al. 2017), community engagement has been an effective model of modifying health behaviours in particular. Hence, community engagement could be an important determinant of success or failure of an intervention aimed at improving immunisation coverage.

Due to the variability in contexts, activities to address the barriers to immunisation through community engagement also vary. Therefore, no single theory of change can capture all the different ways that community engagement will affect immunisation outcomes. Furthermore, there is no strict correspondence between types of engagement and intervention activities. For the most part, each activity can be structured such that community engagement is the intervention (engagement is embedded) or the community is engaged in the design or implementation of the intervention. For example, consider an intervention involving village health committees in which community members are adequately represented. These committees can be created by a sponsoring agency with little or no input from the actual community. In this case, the intervention is engagement. On the other hand, community members may develop a village health committee during the design phase of major clinic infrastructure projects to ensure the efficient management of the intervention. Hence, these committees could arise from community involvement in the design of the intervention. In addition, an existing village health committee could be leveraged for the implementation of community meetings at which committee members lead a dialogue on immunisation.

Broadly, we expect that community engagement activities will increase awareness of vaccine-preventable diseases, knowledge of where and when to get vaccinations, and motivation to get vaccinations among caregivers (Figure 1). Community engagement may also increase skill, motivation, and accountability of health workers. These ought to lead to improved demand for and delivery of services, which will increase the number of vaccinated children and could potentially reduce child morbidity and mortality. It is important to note that child mortality and morbidity are affected by critical factors beyond immunisation, such as access to safe food and adequate nutrition, safe water and
quality care by a trained health provider when needed. Without these critical enabling factors immunisation alone may not be effective in reduction of child mortality and morbidity.
Supplementary Figure 1: Intervention theory of chan
Appendix 3: Search strategy and search terms for impact evaluations

We searched the academic databases and websites listed below on 17 May 2019 and updated the search on 5 May, 2020. A full record of the applied search terms is provided in Appendix C.

1. MEDLINE
2. CAB Global Health
3. EMBASE
4. Cochrane Controlled Trials Register (CENTRAL)
5. CINAHL
6. PsycINFO
7. Pline
8. Africa-wide information
9. Academic search complete
10. Scopus
11. Campbell Library
12. Google Scholar
13. EconLit
14. IDEAS/RePEc
15. WHO Global Index Medicus
16. Pascal-Francis
17. Open-Grey
18. Grey Literature Report
19. Social Science Research Network (SSRN)
20. Eldis
21. GAVI
22. Epistemonikos
23. Innovations for Poverty Action (IPA)
25. 3ie Impact Evaluation Repository
26. 3ie Systematic Review Repository
27. Registry of International Development Impact Evaluations (RIDIE)
28. Global Development Network
29. World Bank Development Impact Evaluation (DIME) and Impact Evaluation Policy Papers
30. Inter-American Development Bank
31. Center for Global Development
32. Center for Effective Global Action (CEGA)
33. DFID Research for Development (R4D)
34. USAID

Website searches for capturing grey literature were carried out from January to May 2020

Search terms

LdMICs

2. (Afghanistan or Albania or Algeria or Angola or Argentina or Armenia or Armenian or Azerbaijan or Bangladesh or Benin or Byelorussian or Belarus or Belorussian or Belarus or Belorussia or Belize or Bhutan or Bolivia or Bosnia or Herzegovina or Hercegovina or Botswana or Brazil or Bulgaria or Burkina Faso or Burkina Faso or Upper Volta or Burundi or Urundi or Cambodia or Khmer Republic or Kampuchea or
Cameroon or Cameroons or Cameroon or Camerons or Cape Verde or Cabo Verde or Central African Republic or Chad or Tchad or China or Colombia or Comoros or Comoro Islands or Comores or Mayotte or Congo or Zaire or Costa Rica or Cote d'Ivoire or Ivory Coast or Cuba or Djibouti or French Somaliland or Dominica or Dominican Republic or East Timor or East Timur or Timor Leste or Ecuador or Egypt or United Arab Republic or El Salvador or Eritrea or Ethiopia or Fiji or Gabon or Gabonese Republic or Gambia or Gaza or Georgia Republic or Georgian Republic or Ghana or Grenada or Guatemala or Guinea or Guinea or Guyana or Haiti or Honduras or India or Maldives or Indonesia or Iran or Iraq or Jamaica or Jordan or Kazakhstan or Kazakh or Kenya or Kiribati or Korea or Kosovo or Kyrgyzstan or Kirghizia or Kyrgyz Republic or Kirghiz or Kirgizistan or Lao PDR or Laos or Lebanon or Lesotho or Basutoland or Liberia or Libya or Macedonia or Madagascar or Malagasy Republic or Malaysia or Malaya or Malay or Sabah or Sarawak or Malawi or Mali or Marshall Islands or Mauritania or Mauritius or Agalega Islands or Mexico or Micronesia or Middle East or Moldova or Moldovia or Moldovian or Mongolia or Montenegro or Morocco or Ifni or Mozambique or Myanmar or Burma or Namibia or Nepal or Netherlands Antilles or Nicaragua or Niger or Nigeria or Muscat or Pakistan or Palau or Palestine or Panama or Paraguay or Peru or Philippines or Philippine or Philippines or Papua New Guinea or Romania or Rumania or Roumania or Rwanda or Ruanda or Saint Lucia or St Lucia or Saint Vincent or St Vincent or Grenadines or Saoa or Samoan Islands or Navigator Island or Navigator Islands or Sao Tome or Senegal or Serbia or Montenegro or Seychelles or Sierra Leone or Sri Lanka or Solomon Islands or Somalia or Sudan or Suriname or Surinam or Swaziland or Eswatini or South Africa or Syria or Tajikistan or Tadjikistan or Tadjik or Tanzania or Thailand or Togo or Toogolese Republic or Tonga or Tunisia or Turkey or Turkmenistan or Turkmen or Uganda or Ukraine or Uzbekistan or Uzbek or Vanuatu or New Hebrides or Venezuela or Vietnam or Viet Nam or West Bank or Yemen or Zimbabwe or Rhodesia)

3. (Africa or Asia or Caribbean or West Indies or South America or Latin America or Central America)
4. ((developing or less* developed or least developed or under-developed or underdeveloped or middle income or low* income or underserved or under served or deprived or poor* or resource limited or resource constrained) adj (countr* or nation? or population? or world or state* or emerging econom* or global south))
5. ((developing or less* developed or least developed or under-developed or underdeveloped or middle income or low* income or resource limited or resource constrained) adj (economy or economies))
6. (low* adj (gdp or gnp or gross domestic or gross national))
7. (low adj3 middle adj3 countr*)
8. (lmic or lmics or third world or lami countr*)
9. transitional countr*
10. 1 OR 2 OR 3 OR 4 OR 5 OR 6 OR 7

Immunisation

1. (immuniz* or immunis* or vaccin* or inoculat* or innoculat* or immunotherap* or prophyla*)

Young children and caregivers

1. (child* or infant* or newborn* or neonat* or prenatal or pre natal or antenatal or ante natal or baby or babies or toddler* or preschool* or parent* or mother* or father* or maternal or paternal or caregiver* or grandparent* or grandmother* or grandfather* or family member*)

Impact evaluation methods
1. (random* or experiment* or (match* adj2 (propensity or coarsened or covariate)) or "propensity score" or "difference in difference"* or "difference-in-difference"* or "differences in difference"* or "differences-in-difference"* or "double difference"* or "quasi-experimental" or "quasi experimental" or "quasi experiment" or (estimator or counterfactual) and evaluation*) or "instrumental variable" or (IV adj2 (estimation or approach)) or regression discontinuity or time series or segment* regression

- Example full search strategy

Below we present a draft of the full search strategy used to search MEDLINE. Note that in MEDLINE’s syntax, terms with strokes (e.g., Immunization/) denote Medical Subject Heading (MeSH) terms, while strings appended with “ti,ab,kw” are searched in the title, abstract, and keyword fields of records in the database.

**Ovid MEDLINE(R) and Epub Ahead of Print, In-Process & Other Non-Indexed Citations, Daily and Versions(R) <1946 to May 15, 2019> Searched 16th May 2019**

1  (immuniz* or immunis* or vaccin* or inoculat* or innoculat* or immunotherap* or prophyla*).ti,ab,kw. (672732)

2  immunization/ or immunization, passive/ or immunization schedule/ or immunization, secondary/ or immunotherapy, active/ or vaccination/ or Immunization Programs/ or mass vaccination/ (158163)

3  Tuberculosis Vaccines/ or BCG Vaccine/ or Diphtheria-Tetanus Vaccine/ or Meningococcal Vaccines/ or Pertussis Vaccine/ or Diphtheria-Tetanus-acellular Pertussis Vaccines/ or Diphtheria-Tetanus-Pertussis Vaccine/ or Diphtheria-Tetanus Vaccine/ or Measles Vaccine/ or Mumps Vaccine/ or Rubella Vaccine/ or Measles-Mumps-Rubella Vaccine/ or Poliovirus Vaccines/ or Poliovirus Vaccine, Inactivated/ or Poliovirus Vaccine, Oral/ or Japanese Encephalitis Vaccines/ or Rotavirus Vaccine/ (49639)

4  or/1-3 (720583)

5  developing countries.sh,kf. (83271)

6  (Africa or Asia or Caribbean or West Indies or South America or Latin America or Central America).ti,ab,kw. (200981)

7  Africa/ or Asia/ or Caribbean/ or West Indies/ or South America/ or Latin America/ or Central America/ (73389)

8  (Africa or Central America or South America or Caribbean or Central Asia or Afghanistan or Albania or Algeria or Angola or Argentina or Armenia or Armenian or Azerbaijan or Bangladesh or Benin or Byelarus or Byelorussian or Belarus or Belorussian or Belorussia or Belize or Bhutan or Bolivia or Bosnia or Herzegovina or Botswana or Brazil or Bulgaria or Burkina Faso or Burkina Fasso or Upper Volta or Burundi or Urundi or Cambodia or Khmer Republic or Kampuchea or Cameron or Cameroon or Camerons or Cape Verde or Cabo Verde or Central African Republic or Chad or Tchad or China or Colombia or Comoros or Comoro Islands or Comores or Mayotte or Congo or Zaire or Costa Rica or Cote d'Ivoire or Ivory Coast or Cuba or Djibouti or French Somaliland or Dominica or Dominican Republic or East Timor or East Timur or Timor Leste or Ecuador or Egypt or United Arab Republic or El Salvador or Eritrea or Ethiopia or Fiji or Gabon or Gabonese Republic or Gambia or Gaza or Georgia Republic or Georgian Republic or Ghana or Grenada or Guatemala or Guinea or Guiana or Guyana or Haiti or Honduras or India or Maldives or Indonesia or Iran or Iraq or Jamaica or Jordan or Kazakhstan or Kazakh or Kenya or Kiribati or Korea or Kosovo or Kyrgyzstan or Kirghizia or Kyrgyz Republic or Kirghiz or Kirgizstan or Lao PDR or Laos or Lebanon or Lesotho or Basutoland or Liberia or Libya or Macedonia or Madagascar or Malagasy Republic or Malaysia or Malay or Sabah or Sarawak or Malawi or Mali or Marshall Islands or Mauritania
or Mauritius or Agalega Islands or Mexico or Micronesia or Middle East or Moldova or Moldovia or Moldovan or Mongolia or Montenegro or Morocco or Ifni or Mozambique or Myanmar or Myanma or Burma or Namibia or Nepal or Netherlands Antilles or Nicaragua or Niger or Nigeria or Muscat or Pakistan or Palau or Palestine or Panama or Paraguay or Peru or Philippines or Philippine or Papua New Guinea or Romania or Rumania or Roumania or Rwanda or Ruanda or Saint Lucia or St Lucia or Saint Vincent or St Vincent or Grenadines or Samoa or Samoan Islands or Navigator Island or Navigator Islands or Sao Tome or Senegal or Serbia or Montenegro or Seychelles or Sierra Leone or Sri Lanka or Solomon Islands or Somalia or Sudan or Suriname or Surinam or Swaziland or Eswatini or South Africa or Syria or Tajikistan or Tadzhikistan or Tajik or Tanza or Tanzania or Thailand or Togo or Togolese Republic or Tonga or Tunisia or Turkey or Turkmenistan or Turkmen or Uganda or Ukraine or Uzbekistan or Uzbek or Vanuatu or New Hebrides or Venezuela or Vietnam or Viet Nam or West Bank or Yemen or Zambia or Zimbabwe or Rhodesia).ti,ab,kw.sh. (1387951)

9  ((developing or less* developed or least developed or under developed or underdeveloped or middle income or low* income or underserved or under served or deprived or poor* or resource limited or resource constrained) adj (countr* or nation? or population? or world or state*)).ti,ab,kw. (91955)

10  ((developing or less* developed or least developed or under developed or underdeveloped or middle income or low* income or resource limited or resource constrained) adj (economy or economies)).ti,ab,kw. (504)

11  (low* adj (gdp or gnp or gross domestic or gross national)).ti,ab,kw. (234)

12  (low adj3 middle adj3 countr*).ti,ab,kw. (12819)

13  (lmic or lmic* or third world or lami countr*).ti,ab,kw. (6486)

14  (transitional countr* or emerging econom* or global south).ti,ab,kw. (873)

15  or/5-14 (1487811)

16  4 and 15 (80257)

17  Parents/ or Fathers/ or Mothers/ or Grandparents/ or Caregivers/ or Single Parent/ or Pregnant Women/ or Child, Preschool/ or Infant/ or Infant, Newborn/ or Infant, Low Birth Weight/ or Infant, Small for Gestational Age/ or Infant, Very Low Birth Weight/ or Infant, Extremely Low Birth Weight/ or Infant, Postmature/ or Infant, Premature/ or Infant, Extremely Premature/ (1597959)

18  (child* or infant* or newborn* or neonat* or neo nat* or prenatal or pre natal or ante natal or antenatal or baby or babies or toddler* or preschool* or parent* or mother* or father* or maternal or paternal).ti,ab,kw. (2284298)

19  or/17-18 (2893768)

20  16 and 19 (28515)

21  (random* or experiment* or (match* adj2 (propensity or coarsened or covariate)) or "propensity score" or ("difference in difference" or "difference-in-difference" or "differences in difference" or "differences-in-difference" or "double difference") or ("quasi-experimental" or "quasi experimental" or "quasi-experiment" or "quasi experiment") or ((estimator or counterfactual) and evaluation*) or "instrumental variable" or (IV adj2 (estimation or approach)) or regression discontinuity or time series or segment* regression).ti,ab,kw. (2983909)
22 Randomized Controlled Trial/ or Random Allocation/ or Evaluation Studies/ or Propensity Score/ or Interrupted Time Series Analysis/ or Controlled Before-After Studies/ or Controlled Clinical Trial/ or Non-Randomized Controlled Trials as Topic/ (896249)

or/21-22 (3415351)

24 Cost Analysis/ or Cost-Benefit Analysis/ or Quality-Adjusted Life Years/ or Economics, Medical/ or Cost of Illness/ or Health Care Costs/ or Direct Service Costs/ or Budgets/ or Health Care Sector/ or Public Expenditures/ (193420)

(cost-effective* or cost-benefit).ti,ab,kw. (130711)

26 ("life year" or "life years" or qaly* or daly*).ti,ab,kw. (18749)

27 ((economic* or cost*) adj6 (mortality or death* or markov)).ti,ab,kw. (17246)

28 ("cost minimi*" or "cost-utiliti*" or "economic evaluation*" or "economic review*" or "cost outcome" or "cost analys*" or "economic analys*" or "budget* impact analys*").ti,ab,kw. (27594)

or/24-28 (301485)

29 (review or meta-analysis).pt. (2557796)

30 meta-analysis/ or "systematic review"/ (163880)

32 cochrane database of systematic reviews.jn. (14162)

33 (systematic review or literature review).ti. (130870)

or/30-33 (2590211)

35 23 or 29 or 34 (5946083)

36 20 and 35 (8309)

37 exp Animals/ (22312530)

38 Humans/ (17732234)

39 37 not (37 and 38) (4580296)

40 36 not 39 (7984)
Appendix 4: List of excluded impact evaluation studies

We excluded the following studies at full-text screening for not meeting our community engagement criteria.

Reference list


A Quasi-Experimental Comparison of Impacts on Health Service Indicators between Hospitals and Health Centers in Malawi.” Health Systems & Reform 6 (1). https://doi.org/10/gg2njg.


Djibuti, Mamuka, George Gotsadze, Akaki Zoidze, George Mataradze, Laura C Esmail, and Jillian Clare Kohler. 2009. “The Role of Supportive Supervision on Immunization Program Outcome - a Randomized Field Trial from Georgia.” BMC International Health and Human Rights 9 (S1). https://doi.org/10/dxd2rk.


Appendix 5: Qualitative and cost-evidence search strategy

Qualitative search protocol

Purpose
The purpose of this review is to inform policy recommendations related to community engagement interventions to improve immunisation outcomes in low- and middle-income countries. Specifically, we would like to provide insights related to (1) what to do / not do and (2) where to do it / not do it. We want to know not only what works, but why and how.

Our primary, qualitative research question is: What factors relating to programme design, implementation, context, and mechanism are associated with better or worse outcomes along the causal chain? Do these vary by the level of community engagement?

Sub questions are:
1. What pre-intervention activities and characteristics of local context facilitate or inhibit the effectiveness of an intervention?
   a. What characteristics of local context facilitate or inhibit the effectiveness of an intervention? These include but are not limited to feasibility, clinic readiness, acceptability, need, barriers, and demographic composition.
   b. What is the evidence regarding the likelihood that pre-intervention activities, including diagnostics, formative research, and community engagement, affect the effectiveness of an intervention?
2. What is the evidence supporting or refuting the existence of certain steps along the ToC?
   a. Which steps in the causal chain tend to be barriers to or facilitators of impact? Where does the causal chain tend to be interrupted?
   b. What time varying contextual factors affect the relationships seen in the ToC?

In order to respond to these questions, 3ie seeks to identify qualitative papers related to the quantitative papers included in the broader systematic review on effectiveness of community engagement interventions for improving immunisation outcomes. This information will contextualise the results of the quantitative papers. The qualitative papers will be instrumental in expanding our understanding of why and how certain interventions were successful (or not).

This document outlines the suggested approach for identifying these qualitative papers.

Required steps
1. Create a unique dropbox folder for each study using the first author’s last name and year as the folder title
2. For each step outlined below, save all potentially relevant documents to this file as they are identified.
   a. Consider this stage to be equivalent to title and abstract screening. If you identify an article that can be quickly discarded as irrelevant (such as budgetary information), there is no need to include. However, if you think that there may be any information related to our research questions, please include at this stage.
   b. If relevant websites are identified that cannot be easily converted to PDF format, create a word document. Write the title of the webpage and provide the link below it.

If a trial registration number for clinicaltrials.gov is provided
Use this approach if a trial registration number for clinicaltrials.gov is provided. This will be referenced in the abstract and / methods section and appear in a form similar to that found in Robertson et al 2013:

This trial is registered with ClinicalTrials.gov, number NCT00966849.

If no trial registration number is indicated, proceed to the following section (Section 2: If the project has a unique name).

1.1 Go to the website clinicaltrials.gov

1.2 Under “find a study” in the field “other terms,” enter the trial registration number.

1.3 There should only be one hit. Click on the title under “study title”

1.4 Three tabs should be available under the basic study information. They read “study details,” “tabular view,” and “study results.” Select “study results”

1.5 Scroll to the bottom of the page. Under “more information” there are sections titled “Publication of results” and “other publications”
   a. In most (but not all) cases, there should be a list of publications here. Save the relevant files to the folder

After this, proceed to section 2.

If the project has an unique name

Use this approach if the intervention is named. If a project name is not given in the paper, but the project is registered on clinicaltrials.gov, use the project title provided there.

2.1 Search the project name on google scholar.
   a. Review each search result until you find 5 in a row that are irrelevant.
   b. At this point, stop and proceed to step 2.2.

2.2 If the funder’s name is not provided, proceed directly to step 2.3. If the funder’s name is provided, search the project name on the funder’s website.
   a. Review each search result until you find 5 in a row that are irrelevant.
   b. At this point, stop and proceed to step 2.3.

2.3 If the implementer’s name is not provided, proceed directly to section 3. If the implementer’s name is provided, search the project name on the implementer’s website.
   a. Review each search result until you find 5 in a row that are irrelevant.

Other options

If the trial is not registered and does not have a formal name, the following approaches can be used as a last resort:
3.1 If funder and / or implementer names are provided, go to their websites. Search for the article on their website. It may link to a program page.

   a. If no program page is identified, try to manually search through their website.

3.2 Search the full article name on google and google scholar. Review the top 10 hits on each to see if they provide more information about the program.

If during either of these steps, a program name and / or trial registration number are identified, return to the relevant steps above.

**Searching funder and implementer websites**

Some of these websites will have good, built-in search functions. If these are useful, use them. However, if the provided search functionality is not adequate, type site:HomepageOfFunderWebsite SearchTerm. For example, if I wanted to search the World Bank website for information about COVID, I would search “site:www.worldbank.org COVID” in google (Note: no space between “site:” and the url). This will provide me hits from the World Bank website related to COVID.

**If no linked articles are identified**

At the end of each day, email Avantika (abagai@3ieimpact.org) and Charlotte (clane@3ieimpact.org) the title of all papers for which you were unable to identify linked articles.

**Cost evidence search protocol**

The evaluations included in the cost and cost-effectiveness evidence were drawn from the 61 evaluations that were identified for the systematic review. Since cost reporting and analysis is often not required by donors or included in impact evaluations of global development interventions, we undertook structured outreach to the authors of the 61 included studies to request any additional cost evidence. In case the lead or corresponding author email addresses were nonfunctional, we contacted one of the co-authors.

We successfully contacted 58 out of 61 study authors that reported intervention effectiveness only or intervention effectiveness and partial cost information and requested estimates of the intervention costs as well as any additional raw data files, summary tables, sensitivity analyses, cost evidence, published reports or notes, and non-published documentation, methods or analyses relating to cost-effectiveness. In the first two weeks of this request, our team obtained responses from a total of 25 study authors, of which 12 produced or pointed us in the direction of economic evidence for their respective interventions and 7 authors confirmed that any such costing analysis was not undertaken as part of their study. We excluded any studies that did not report estimates of both costs and effectiveness and identified 22 evaluations with both cost and effectiveness estimates.

**List of included qualitative studies**


Grayman, Jesse Hession, Siti Ruhanawati, and Nelti Anggraini. 2014. "Opportunities and approaches for better nutrition outcomes through PNPM generasi : a qualitative study."


Utilization of Maternal and Child Health Services and Basic Education Services in the Provinces of West Java and East Nusa Tenggara. In SMERU Research Institute: SMERU Research Institute.


**List of included other studies**


Adamu, Abdu Abdullahi. 2019. "Using quality improvement approach to address missed opportunities for vaccination in Kano Metropolis, Nigeria." PhD, Faculty of Medicine and Health Sciences, Stellenbosch University.

Alhassan, Robert Kaba. 2017. "Healthcare quality in Ghana: Improving healthcare quality and health worker motivation to promote sustainable health insurance." PhD, Faculty of Medicine (AMC-UvA), University of Amsterdam.


Borkum, Evan.; Anitha; Sivasankaran, Divya; Vohra, Swetha; Sridharan, Dana; Rotz, Lakshmi; Ramnkrishnan, Mercy; Manoranjini, Sukhmani; Sethi, and Anu. Rangarajan. 2015. "Evaluation of the Team-Based Goals and Performance-Based Incentives (TBGI) Intervention in Bihar: Findings from the 30-month Follow-up." *Mathematica Policy Research.*


Mitchell, Steven, Neil Andersson, Noor Mohammad Ansari, Khalid Omer, José Legorreta Soberanis, and Anne Cockcroft. 2009. "Equity and vaccine uptake: a cross-sectional study of measles vaccination in Lasbela District,


Padayachee, Thesandree;, Natasha; Chetty, Muzi; Matse, Tumelo; Mampe, and Helen Schneider. 2013. Progress in the establishment of ward based outreach teams: experiences in the North West Province. University of Western Cape.


Robertson, Laura, Phyllis Mushati, Jeffrey W. Eaton, Lorraine Sherr, Jeremiah C. Makoni, Morten Skovdal, Tom Crea, Gideon Mavise, Lovemore Dumba, Christina Schumacher, Shungu Munyati, Constance Nyamukapa, and Simon Gregson. 2012. "Household-based cash transfer targeting strategies in Zimbabwe: are we reaching the


UNICEF. 2010. Evaluation report of the community health strategy implementation in Kenya. Division of Community Health Services


Appendix 6: Data extraction tools

Quantitative data extraction tool

<table>
<thead>
<tr>
<th>VARIABLE LABEL</th>
<th>EXPLANATION</th>
</tr>
</thead>
<tbody>
<tr>
<td>Study ID</td>
<td>This is the study ID - it should match the study ID from the Outcome Mapping Sheet (e.g., SC-SR_1)</td>
</tr>
<tr>
<td>Estimate ID</td>
<td>The estimate ID will provide a specific number for each effect size extracted and should include the original study number, underscore, then the unique ID number (e.g., SC-SR1_1, SC-SR1_2 and so on)</td>
</tr>
<tr>
<td>Author</td>
<td>For 1 author: leading author last name (e.g., Gomez)  For 2 authors: both author last names with ampersand in between (e.g., Smith &amp; Bahn)  For 3 or more authors: leading author last name followed by et al. (e.g., Gupta et al.)</td>
</tr>
<tr>
<td>Year</td>
<td>Year published</td>
</tr>
<tr>
<td>Design</td>
<td>0=Experimental Design (e.g., RCT), 1=Quasi-Experimental Design</td>
</tr>
<tr>
<td>How Counterfactual is Chosen?</td>
<td>Free text (e.g., random control trial, propensity score matching, etc) - Multiple codes are ok</td>
</tr>
<tr>
<td>Analysis type for this effect size</td>
<td>Free text, what type of analysis was used (Regression, 2SLS, ANCOVA, etc.)- Multiple codes are ok</td>
</tr>
<tr>
<td>Country</td>
<td>Country of intervention</td>
</tr>
<tr>
<td>Region</td>
<td>Region/continent of intervention</td>
</tr>
<tr>
<td>Estimate Type</td>
<td>Type of data for this effect size: 1 = Continuous - means and SDs, 2 = Continuous - mean difference and SD, 3 = Dichotomous outcome - proportions, 4 = Regression data - dichotomous outcome (e.g., logistic regression) 5 = Regression data - continuous outcome (e.g., linear regression)</td>
</tr>
<tr>
<td>Comparison</td>
<td>1=No intervention (service delivery as usual), 2=Other intervention, 3=Pipeline (wait-list) control (still service delivery as usual)</td>
</tr>
<tr>
<td>Describe Comparison Group</td>
<td>If answer above is (1) no intervention, type N/A, if (2) Other Intervention, list what intervention the control group is receiving, if (3) Pipeline control, report when the control group will receive the intervention in relation to the treatment group (e.g., one year later)</td>
</tr>
<tr>
<td>Subgroup</td>
<td>Is this analysis of a subgroup? 0=no, 1=yes</td>
</tr>
<tr>
<td>If yes to subgroup, describe</td>
<td>Free text, describe the subgroup if applicable (e.g., boys, girls). If no subgroup, type N/A</td>
</tr>
<tr>
<td>-------------------------------</td>
<td>----------------------------------------------------------------------------------</td>
</tr>
<tr>
<td>Source</td>
<td>Note the page number, table number, column, and row you used to extract the data</td>
</tr>
<tr>
<td>Treatment Effect</td>
<td>1=Intention to Treat (ITT), 2=Average Treatment Effect on the Treated (ATET), 3=Average Treatment Effect (ATE) 4 = Local Average Treatment Effect (LATE)</td>
</tr>
<tr>
<td>Intervention</td>
<td>Free text, what is the intervention</td>
</tr>
<tr>
<td>Engagement as intervention</td>
<td>1=yes, 0=No</td>
</tr>
<tr>
<td>Engagement as intervention: developing community buy-in</td>
<td>1=yes, 0=No</td>
</tr>
<tr>
<td>Engagement as intervention: creation of new cadres or health committees</td>
<td>1=yes, 0=No</td>
</tr>
<tr>
<td>Engagement as intervention: both</td>
<td>1=yes, 0=No</td>
</tr>
<tr>
<td>Engagement in design</td>
<td>1=yes, 0=No</td>
</tr>
<tr>
<td>Engagement in design: community decision making</td>
<td>1=yes, 0=No</td>
</tr>
<tr>
<td>Engagement in design: community feedback</td>
<td>1=yes, 0=No</td>
</tr>
<tr>
<td>Engagement in implementation</td>
<td>1=yes, 0=No</td>
</tr>
<tr>
<td>Engagement in implementation: governance and decisions</td>
<td>1=yes, 0=No</td>
</tr>
<tr>
<td>Engagement in implementation: provision of resources</td>
<td>1=yes, 0=No</td>
</tr>
<tr>
<td>Multiple engagement types</td>
<td>1=yes, 0=No</td>
</tr>
<tr>
<td>Engagement in design and engagement as intervention</td>
<td>1=yes, 0=No</td>
</tr>
<tr>
<td>Engagement in design and implementation</td>
<td>1=yes, 0=No</td>
</tr>
<tr>
<td>Engagement in implementation and engagement as intervention</td>
<td>1=yes, 0=No</td>
</tr>
<tr>
<td>Exposure to intervention (in months)</td>
<td>How long is the intervention exposure itself?</td>
</tr>
<tr>
<td>Evaluation period (in months)</td>
<td>The total number of months elapsed between offering an intervention and the point at which an outcome measure is taken post intervention, or as a follow-up measurement. If less than one month, use decimals (e.g., one week would be .25)</td>
</tr>
<tr>
<td>Post-intervention or change from baseline?</td>
<td>0 = Post-intervention, 1 = Change from baseline</td>
</tr>
<tr>
<td>------------------------------------------</td>
<td>------------------------------------------------</td>
</tr>
<tr>
<td>Source of the Outcome data (only use for outcomes in category JAA)</td>
<td>1=Immunisation Card, 2=Recall, 3=Combination of both immunisation card and recall, 4=Health Admin Data, 5=N/A</td>
</tr>
<tr>
<td>Author definition of outcome</td>
<td>Free text - How does the author define the outcome?</td>
</tr>
</tbody>
</table>

**OUTCOME CODES**

**Thinking and Feeling**
Code 1 under any applicable columns. See OUTCOME GUIDANCE document for further explanations.

**Social Processes**
Code 1 under any applicable columns. See OUTCOME GUIDANCE document for further explanations.

**Readiness to Vaccinate**
Code 1 under any applicable columns. See OUTCOME GUIDANCE document for further explanations.

**Practical Factors**
Code 1 under any applicable columns. See OUTCOME GUIDANCE document for further explanations.

**Community Health Workers**
Code 1 under any applicable columns. See OUTCOME GUIDANCE document for further explanations.

**Vaccinators**
Code 1 under any applicable columns. See OUTCOME GUIDANCE document for further explanations.

**IAC Admins**
Code 1 under any applicable columns. See OUTCOME GUIDANCE document for further explanations.

**Health information systems**
Code 1 under any applicable columns. See OUTCOME GUIDANCE document for further explanations.

**Vaccine Availability**
Code 1 under any applicable columns. See OUTCOME GUIDANCE document for further explanations.

**Resources**
Code 1 under any applicable columns. See OUTCOME GUIDANCE document for further explanations.

**Vaccination coverage**
Code 1 under any applicable columns. See OUTCOME GUIDANCE document for further explanations.

**Health Outcomes**
Code 1 under any applicable columns. See OUTCOME GUIDANCE document for further explanations.

**EFFECT SIZE DATA EXTRACTION**

**Reverse Sign (i.e., decrease is good)**
Record 0='no' if an increase is good, record 1='yes' if a decrease is good and the sign needs to be reversed.

**Unit of analysis**
What is the unit of analysis? UOA for this effect size: 1= Individual, 2= Household, 3= Group (e.g. community
<p>| <strong>mean_t</strong> | Outcome mean for the treatment group |
| <strong>sd_t</strong> | Outcome standard deviation for treatment group |
| <strong>mean_c</strong> | Outcome mean for the comparison group |
| <strong>sd_c</strong> | Outcome standard deviation for control group |
| <strong>mean_overall_diff</strong> | Overall mean difference (treatment - control) |
| <strong>diff se</strong> | Standard error of the overall mean difference |
| <strong>Diff_t</strong> | t-statistic of mean difference |
| <strong>Odds ratio</strong> | Odds ratio reported in the study |
| <strong>OR_se</strong> | Odds ratio standard error reported in the study |
| <strong>Risk ratio</strong> | Risk ratio reported in study |
| <strong>RR_se</strong> | Risk ratio standard error |
| <strong>reg_coeff</strong> | Report the regression coefficient of the treatment effect |
| <strong>reg_SE</strong> | Report the associated standard error of the regression coefficient. |
| <strong>reg_t</strong> | Report the associated t statistic of the effect size (coefficient/SE) |
| <strong>Exact p value</strong> | Exact p value if given, if not, record as written in the manuscript (e.g., p &lt; .001, or p &gt; .05) |
| <strong>clust_t</strong> | Number of clusters - treatment group |
| <strong>clust_c</strong> | Number of clusters - control group |
| <strong>clust_T</strong> | Number of clusters - total sample |
| <strong>n_t</strong> | Sample size - treatment group |
| <strong>n_c</strong> | Sample size - control group |
| <strong>n_T</strong> | Sample size - total sample |
| <strong>Does the sample size need to be adjusted (if so complete column CL)?</strong> | Code as 'yes' or no |
| <strong>periods (1 if cross sectional)</strong> | Record how many periods of evaluation there are (e.g., cross section is 1, panel data with 3 measurements is 3) |
| <strong>Treatment Variable</strong> | Record the treatment variable as written in the model (e.g., the variable name the author uses, such as &quot;Intervention x Time&quot;) |
| <strong>dataset</strong> | Record if data comes from an identified dataset |</p>
<table>
<thead>
<tr>
<th>coder</th>
<th>Record your name</th>
</tr>
</thead>
<tbody>
<tr>
<td>Notes</td>
<td>Record any notes important for the team</td>
</tr>
<tr>
<td>n_T_revised</td>
<td>THIS IS FOR SENIOR QUANT LEAD TO FILL OUT</td>
</tr>
<tr>
<td>sp</td>
<td>THIS IS FOR SENIOR QUANT LEAD TO FILL OUT</td>
</tr>
<tr>
<td>d</td>
<td>THIS IS FOR SENIOR QUANT LEAD TO FILL OUT</td>
</tr>
<tr>
<td>g</td>
<td>THIS IS FOR SENIOR QUANT LEAD TO FILL OUT</td>
</tr>
<tr>
<td>var(d)</td>
<td>THIS IS FOR SENIOR QUANT LEAD TO FILL OUT</td>
</tr>
<tr>
<td>se(d)</td>
<td>THIS IS FOR SENIOR QUANT LEAD TO FILL OUT</td>
</tr>
<tr>
<td>CI_l</td>
<td>THIS IS FOR SENIOR QUANT LEAD TO FILL OUT</td>
</tr>
<tr>
<td>CI_u</td>
<td>THIS IS FOR SENIOR QUANT LEAD TO FILL OUT</td>
</tr>
<tr>
<td>remove</td>
<td>THIS IS FOR PROJECT MANAGER TO FILL OUT</td>
</tr>
<tr>
<td>Formula Used</td>
<td>THIS IS FOR SENIOR QUANT LEAD TO FILL OUT</td>
</tr>
<tr>
<td>yi_1</td>
<td>THIS IS FOR SENIOR QUANT LEAD TO FILL OUT</td>
</tr>
<tr>
<td>yi_rev</td>
<td>THIS IS FOR SENIOR QUANT LEAD TO FILL OUT</td>
</tr>
<tr>
<td>yi</td>
<td>THIS IS FOR SENIOR QUANT LEAD TO FILL OUT</td>
</tr>
<tr>
<td>vi</td>
<td>THIS IS FOR SENIOR QUANT LEAD TO FILL OUT</td>
</tr>
<tr>
<td>wi</td>
<td>THIS IS FOR SENIOR QUANT LEAD TO FILL OUT</td>
</tr>
<tr>
<td>ywi</td>
<td>THIS IS FOR SENIOR QUANT LEAD TO FILL OUT</td>
</tr>
<tr>
<td>95ci_lower</td>
<td>THIS IS FOR SENIOR QUANT LEAD TO FILL OUT</td>
</tr>
<tr>
<td>95ci_upper</td>
<td>THIS IS FOR SENIOR QUANT LEAD TO FILL OUT</td>
</tr>
<tr>
<td>cilow_3sf</td>
<td>THIS IS FOR SENIOR QUANT LEAD TO FILL OUT</td>
</tr>
<tr>
<td>cihigh_3sf</td>
<td>THIS IS FOR SENIOR QUANT LEAD TO FILL OUT</td>
</tr>
<tr>
<td>ci</td>
<td>THIS IS FOR SENIOR QUANT LEAD TO FILL OUT</td>
</tr>
<tr>
<td>wb_yi</td>
<td>THIS IS FOR SENIOR QUANT LEAD TO FILL OUT</td>
</tr>
<tr>
<td>Checked</td>
<td>THIS IS FOR EFFECT SIZE RELIABILITY CHECKER TO FILL OUT</td>
</tr>
<tr>
<td>ROB Category</td>
<td>THIS IS FOR SENIOR QUANT LEAD OR PM TO FILL OUT</td>
</tr>
</tbody>
</table>
### Qualitative data extraction tool

<table>
<thead>
<tr>
<th>Name</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Causal mechanism</strong></td>
<td>This set of codes describes the causal mechanisms observed or hypothesized</td>
</tr>
<tr>
<td><strong>Barriers to immunisation</strong></td>
<td>Select any state reasons that people were not immunised. Try to use previously added codes to reduce possible redundancies, but add codes if needed.</td>
</tr>
<tr>
<td><strong>Demand side barriers</strong></td>
<td>Barriers related to low demand for immunisations</td>
</tr>
<tr>
<td><strong>Fear</strong></td>
<td>Low demand for immunisation due to fear of side effects or other factors</td>
</tr>
<tr>
<td><strong>Financial</strong></td>
<td>Select the financial reason that people were not immunised. Add sub-codes as needed.</td>
</tr>
<tr>
<td><strong>Cost of immunisation</strong></td>
<td>People were not immunised due to the cost of the immunisation. Do not add sub-codes.</td>
</tr>
<tr>
<td><strong>Cost of transport</strong></td>
<td>People were not immunised due to the cost of transport.</td>
</tr>
<tr>
<td><strong>Opportunity costs</strong></td>
<td>People were not immunised due to opportunity costs</td>
</tr>
<tr>
<td><strong>Knowledge of Immunisation</strong></td>
<td>Low demand for immunisation due to limited knowledge of its importance. This includes instances when the person simply does not feel that immunisations are important (regardless of &quot;book knowledge&quot;)</td>
</tr>
<tr>
<td><strong>Schedule</strong></td>
<td></td>
</tr>
<tr>
<td><strong>Logistics or distance</strong></td>
<td>Barriers related to logistics or distance. For example, it is too far away, I need to arrange for XXX to happen while I am gone...</td>
</tr>
<tr>
<td><strong>Self-efficacy</strong></td>
<td>Issues related to not feeling empowered to get vaccinated. This includes not knowing how to go about getting vaccinations. For example, not knowing when or where to get vaccinated</td>
</tr>
<tr>
<td><strong>Social norms</strong></td>
<td>Low demand for immunisation due to social pressures and social norms</td>
</tr>
<tr>
<td><strong>Understanding of importance</strong></td>
<td>Low demand for immunisation due to limited knowledge of its importance. This includes instances when the person simply does not feel that immunisations are important (regardless of &quot;book knowledge&quot;)</td>
</tr>
<tr>
<td><strong>Supply side barriers</strong></td>
<td>Limitations to supply that cause problems for immunisation</td>
</tr>
<tr>
<td><strong>Availability of services</strong></td>
<td>Services may not be physically present. This includes clinics being closed.</td>
</tr>
<tr>
<td><strong>Human resources</strong></td>
<td>Limited human resources may reduce the availability of immunisations. Ex: Appropriate staff were not present to provide the injection</td>
</tr>
<tr>
<td><strong>Lack of Infrastructure or supplies</strong></td>
<td>Use code if the lack of immunisation was because there were no drugs, vaccines, cold chain or other items related to physical infrastructure of the health facility</td>
</tr>
<tr>
<td>Poor quality services</td>
<td>Immunisations are not common because there is low quality of service. Providers may not be cooperative, slow, unprofessional/rude, unempathetic, abusive, or inconsistent.</td>
</tr>
<tr>
<td>------------------------------------------</td>
<td>--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------</td>
</tr>
<tr>
<td>Lack of accountability</td>
<td>Use code for text referring to issues related to accountability, corruption, governance and transparency in providing services.</td>
</tr>
<tr>
<td>Lack of Motivation</td>
<td>Use if for text indicating that the health workers are not motivated to perform their duties due to any reason like lack of monetary incentives, recognition, low job satisfaction and other issues.</td>
</tr>
<tr>
<td>Baseline conditions and descriptive characteristics</td>
<td>Discussions of the baseline conditions and descriptive characteristics of the region or population in which the intervention took place. This does not necessarily come from a baseline survey, but can be through other sources. Do not add sub-codes</td>
</tr>
<tr>
<td>Health systems context</td>
<td>Discussion of the health system functioning and context Example: The clinic served a population of 1,000 people Do not add sub-codes</td>
</tr>
<tr>
<td>Immunisation rates</td>
<td>Discussion of baseline immunisation rates. Example: Immunisation rates were low at baseline. Do not add sub-codes</td>
</tr>
<tr>
<td>Political context</td>
<td>Discussion of the political context Example: Political unrest made the intervention difficult to implement. Do not add sub-codes</td>
</tr>
<tr>
<td>Socioeconomic status and demographics</td>
<td>Discussions of SES and demographics at baseline. Example: The community was mostly farmers. Population size Do not add sub-codes</td>
</tr>
<tr>
<td>Facilitators of immunisation</td>
<td>Select any state reasons that people were immunised. Try to use previously added codes to reduce possible redundancies, but add codes if needed.</td>
</tr>
<tr>
<td>Demand Side Facilitators</td>
<td></td>
</tr>
<tr>
<td>Favourable population SES characteristics</td>
<td>For instance those from a higher wealth quantile, urban areas or mothers with higher educational qualifications are more likely to get their children immunised. Do not add sub-codes.</td>
</tr>
<tr>
<td>Knowledge of immunisation schedule</td>
<td>Use this code if caregivers are aware of all immunisations in the schedule and know which vaccination is due when. Do not add sub-codes.</td>
</tr>
<tr>
<td>Lack of fear</td>
<td>Text indicating that there was no fear or vaccine hesitancy. Do not add sub-codes</td>
</tr>
<tr>
<td>Lack of financial constraints</td>
<td>Caregivers can afford the cost of immunisation, cost of transportation to visit healthcare facilities, taking children for immunisation does not come at a high opportunity cost for the caregiver. Do not add sub-codes.</td>
</tr>
<tr>
<td>Lack of logistics or distance related challenges</td>
<td>For instance, caregivers were able to access health services because they lived within a 5km radius of a healthcare facility or they received outreach services regularly. Do not add sub-codes.</td>
</tr>
<tr>
<td>Self-efficacy</td>
<td>Caregivers feel empowered to get their children vaccinated. For example, they know when and where to get vaccinated. Do not add sub-codes.</td>
</tr>
<tr>
<td>Social norms</td>
<td>The social environment owing to prevailing social pressures and norms is encouraging towards child immunisation. For instance, the mother in law in the household or the traditional/religious leaders in the community support immunisation of children. Do not add sub-codes.</td>
</tr>
<tr>
<td>Understanding of importance</td>
<td>Caregivers are aware about the importance of immunisation and are willing to get their children vaccinated. For example, caregivers are able to correctly identify vaccine preventable illness and therefore, more likely to get their children immunised. Do not add sub-codes.</td>
</tr>
</tbody>
</table>

**Supply-side facilitators**

| Availability of infrastructure or supplies | Good quality of services |
| | High accountability |
| | High motivation |
| Human resources | |

| Availability of services | Impacts |
| | What impacts were ultimately achieved? Use this code for general impacts evaluated through the IE that do not fall into one of the other groups. Do not add sub-codes |

| Descriptive impacts | Include authors’ descriptions of impacts that were not directly measured through the IE. Ex: There was generally an increase over time in most sites |

| Equity considerations | If impacts on marginalized groups are considered, select the type of marginalized group considered. Do not add sub-codes |

| Ethnic minority | Impacts on ethnic minorities Do not add sub-codes |

| Other | Impacts on other marginalized groups, like scheduled castes in India. Do not add sub-codes |

| Religious minority | Impacts on religious minorities Do not add sub-codes |

| Women | Impacts on women Do not add sub-codes |
### Long term impacts
Impact measured over a timeframe of more than 5 years from the end of the intervention  Do not add additional sub-codes

### Unintended impacts
Only use this sub-code if the impact is stated as unintended. Do not add sub-codes

### Reasons for project success or failure
Select text in which authors discuss the reasons for the success or failure of the project. This could be related to intervention design, implementation, or the research approach. Example: The intervention was not successful due to a lack of political will. Do not add sub-codes

### Failure
**Implementation or scale-up challenges**
- Competing priorities of health workers
- Other delays, disruptions or implementation variability
- Payment delays

**Intervention**
- Design - Engagement
- Design - other
- Duration, frequency and exposure or reach
- Not accounting for existing constraints or uncontrollable contextual trends

- Caregiver competing priorities
- Civil unrest or political instability or natural calamity
- Fear of AEFI

**Health system issues**
- Health service access or availability
- Poor quality of service or infrastructure

Also includes issues related to health worker demotivation, recruitment or retention.
<table>
<thead>
<tr>
<th>High baseline coverage</th>
<th>The interventions are not geared to/design to achieve last mile coverage or breakthrough stagnating rates.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Lack of awareness or understanding</td>
<td></td>
</tr>
<tr>
<td>Migration</td>
<td></td>
</tr>
<tr>
<td>Resource constraints or scarcity</td>
<td>Infrastructure, electricity, cellular network, etc.</td>
</tr>
<tr>
<td>Social norms</td>
<td></td>
</tr>
<tr>
<td>Wider socio-economic or health-related progress</td>
<td></td>
</tr>
<tr>
<td>Study design</td>
<td></td>
</tr>
<tr>
<td>Contamination</td>
<td></td>
</tr>
<tr>
<td>Other methodological shortcomings</td>
<td></td>
</tr>
<tr>
<td>Mixed results</td>
<td></td>
</tr>
<tr>
<td>Success</td>
<td></td>
</tr>
<tr>
<td>Existing or changing favourable factors or contextual trends</td>
<td></td>
</tr>
<tr>
<td>Health system enablers</td>
<td></td>
</tr>
<tr>
<td>Good quality services and infrastructure</td>
<td>Includes enablers like good outreach, service monitoring, etc.</td>
</tr>
<tr>
<td>Health service access and availability</td>
<td></td>
</tr>
<tr>
<td>Self efficacy, social norms and awareness</td>
<td></td>
</tr>
<tr>
<td>Socio-economic factors</td>
<td></td>
</tr>
<tr>
<td>Implementation improvements</td>
<td></td>
</tr>
<tr>
<td>Intervention features</td>
<td></td>
</tr>
<tr>
<td>Community engagement</td>
<td></td>
</tr>
<tr>
<td>Behaviour change communication</td>
<td></td>
</tr>
<tr>
<td>Community dialogues</td>
<td></td>
</tr>
<tr>
<td>Community involvement in planning and implementation</td>
<td></td>
</tr>
<tr>
<td>Incentives</td>
<td></td>
</tr>
<tr>
<td>Needs assessments, pilots or stakeholder consultations</td>
<td></td>
</tr>
<tr>
<td>Other types of engagement</td>
<td></td>
</tr>
<tr>
<td>Customisation to local context</td>
<td></td>
</tr>
<tr>
<td>Health system integration and organisational structure</td>
<td></td>
</tr>
<tr>
<td>Health worker training</td>
<td></td>
</tr>
<tr>
<td>Intervention duration, dose and exposure</td>
<td></td>
</tr>
<tr>
<td>Leadership and supportive supervision</td>
<td></td>
</tr>
<tr>
<td>Other design features</td>
<td></td>
</tr>
<tr>
<td>Positive participant or beneficiary views of intervention</td>
<td></td>
</tr>
<tr>
<td>Theory of change</td>
<td></td>
</tr>
</tbody>
</table>

Information related to the ToC. Do not code here, but use sub-codes. Do not add sub-codes.

Assumptions

when discussing a causal chain mechanism, we often make assumptions about what will happen. These assumptions are often implicit. We may print materials with the assumption that people will be able to read them. We may train people with the assumption they speak our language. When these assumptions are stated, use the sub-codes to indicate if these assumptions were validated or not. Do not add sub-codes

Incorrect assumption

This indicates that there was an assumption in the causal chain mechanism and the assumption was not correct. For example, a study assumed that increasing supply would increase immunisations. Here, we are assuming that people want immunisations and that low supply is the problem. If instead, we find that people did not want the immunisations to begin with, the assumption was incorrect and would be indicated here. Do not add sub-codes
Valid assumption

This indicates that there was an assumption in the causal chain mechanism and the assumption was valid. For example, a study assumed that increasing supply would increase immunisations. Here, we are assuming that people want immunisations and that low supply is the problem. If this proves to be true, the text should be selected here. Do not add sub-codes.

Causal chain mechanisms

Descriptions of how change is expected to occur. These causal chains are often represented as flow charts, or at least they can be. Text that reflects a causal chain mechanism will discuss expected actions and reactions. “We did this and we expected that.” Do not add sub-codes.

Conclusions

Authors’ conclusions. Any summary of the take-home message of the article. Tends to be “this worked,” “this did not work,” “in conclusion.” Do not add sub-codes.

Cost

Discussions of costs, including the costs of the intervention and any form of cost-analysis. Only select key “take-home message.” Do not add sub-codes.

Data source

For anything related to causal mechanisms, uptake, and uptake and fidelity challenges, indicate what data source provided this information. Do not add additional sub-codes.

Author notes or experiences

FGDs

Implementer interpretations or experiences

Individual interviews

Literature Review

Use this code if the text alludes to existing evidence on the subject.

Observations

If the data comes from direct observations. This must be enumerators looking at and noting the state of something. If it is simply “general impressions,” use the notes/interpretations or experience codes. Do not add sub-codes.

Intervention description

This set of codes describes the intervention.

Beneficiary selection

Select relevant information regarding how beneficiaries were selected for the intervention. Note: This is how beneficiaries of the intervention were selected, NOT a research sampling design or how research participants were selected. Do not add sub-codes.

Community involvement

Description of how the community was involved in beneficiary selection. Example: A community council decided who was eligible. Do not add sub-codes.

Equity considerations

Descriptions of efforts to ensure beneficiary selection was fair and equitable. Example: There was a quota to ensure adequate representation for ethnic minorities. Do not add sub-groups.
**Other selection**
Other descriptions of the selection process and eligibility criteria. Do not add sub-codes.

**Researcher involvement**
Description of the researcher involvement in selection, including selection criteria designed for academic analysis. Example: If only those who were involved in and RCT received the intervention, then the sampling frame for the RCT would also be how beneficiaries for the intervention were selected. Do not add sub-codes.

**Community engagement**
This set of codes describes community engagement in the interventions.

**Community mobilization**
Select the relevant reference to community mobilization, regardless of how this phrase is used. Do not add sub-codes.

- **Absent**
  Statements that there was no effort at community mobilisation. Example: A challenge was that community members were not aware of the intervention activities. Do not add sub-codes.

- **Definitions of community mobilization**
  The author's definition for community mobilisation. Do not add sub-codes.

- **Present - with trust building**
  Descriptions of community engagement that was done in such a way as to establish trust and buy-in. Example: A local organisation was recruited to inform community members of the intervention and explain its importance. Note: expect significant double coding with "Developing community buy-in, Present". Do not add sub-codes.

- **Present - without trust building**
  Descriptions of community mobilization that was done without establishing trust and buy-in. Example: Simply informing the community of an intervention. Do not add sub-codes.

**Inclusion criteria**
Select the relevant reason this intervention/activity was included in the SR. Inclusion criteria into the research project should fall under "sampling frame." Do not use main code, only use sub-codes. Do not add sub-codes.

**Community engagement as the intervention**
Select the way in which community engagement was used as the intervention itself. Do not add sub-codes.

- **Developing community buy-in**
  Descriptions of activities whose primary purpose was the establishment of community buy-in/trust. Do not add sub-codes.

- **Absent**
  Explicit statements that efforts to establish community buy-in/trust were not made should be assigned this code. Example: Authors state that the intervention proceeded without engaging the community. Do not add sub-codes.

- **Present**
  If there were activities whose primary purpose was to establish community buy-in/trust, select this code. Example: Meetings with local leaders to get their support. Note: expect significant double coding with "community mobilization, present - with trust building." Do not add sub-codes.

- **New cadres**
  Description of the development of new cadres of community based structures or systems for health outreach. Examples: Developing community health workers, health volunteers, or the establishment of committees. Do not add sub-codes.
<table>
<thead>
<tr>
<th>Community engagement in designing the intervention</th>
<th>Select the type of community engagement that was sought before on the design of the intervention. Do not add sub-codes</th>
</tr>
</thead>
<tbody>
<tr>
<td>Absent</td>
<td>The community was not given the opportunity to provide feedback or make decisions on the design of the intervention. Example: Explicit statements that the community was not consulted on the design of the intervention. Do not add sub-codes</td>
</tr>
<tr>
<td>Community decision making</td>
<td>The community was given the opportunity to make decisions regarding the design of the intervention. This could vary from taking iterative feedback from the community to consensus building on the design of the intervention to the design being community led.</td>
</tr>
<tr>
<td>Absent</td>
<td>The community did not have decision making power. Example: The intervention design was established a priori and not subject to change. Do not add sub-codes</td>
</tr>
<tr>
<td>Community led</td>
<td>The community had the ultimate decision-making power and/or the intervention design was community led. Example: The community identified which barriers would be targeted and developed an action plan. Do not add sub-codes.</td>
</tr>
<tr>
<td>Partial</td>
<td>The community had some decision-making power. Example: Decisions were made through collaboration between the community, implementers, and/or researchers. Do not add sub-codes</td>
</tr>
<tr>
<td>Community feedback</td>
<td>The community was given the opportunity to provide feedback on the design of the intervention.</td>
</tr>
<tr>
<td>Formative evaluation or stakeholder consultation</td>
<td>There was a previous pilot in which community feedback was taken. Do not add sub-codes.</td>
</tr>
<tr>
<td>Needs assessment</td>
<td>The present study is a pilot in which community feedback is taken.</td>
</tr>
<tr>
<td>Pilot</td>
<td>There was a pilot in which it is explicitly stated that community feedback was taken. Do not add sub-codes.</td>
</tr>
<tr>
<td>Community engagement in implementation</td>
<td>Communities had some opportunity to affect the implementation of the intervention. For example, community is required to spend resources on building health infrastructure and decides whether the intervention will be implemented in their community or those where community members come together to form governance structures such as health committees. OR The community has some responsibility for making decisions regarding the delivery of an intervention. Do not add sub-codes.</td>
</tr>
<tr>
<td>Governance and decisions</td>
<td></td>
</tr>
<tr>
<td>-------------------------</td>
<td>---</td>
</tr>
<tr>
<td>Absent</td>
<td>Explicit statements that the community did not have decision making power. Example: Statements that the community was excluded from implementation. This may be stated as a problem during the discussion. Do not add sub-codes</td>
</tr>
<tr>
<td>Present</td>
<td>The community has some form of decision making in the implementation of the intervention. This could include a community led governance structure. Example: There was a community group monitoring the implementation of the intervention. Do not add sub-codes</td>
</tr>
<tr>
<td>Provision of resources</td>
<td>The community provides resources during the implementation. Examples: Community provides building materials. Do not add sub-codes</td>
</tr>
</tbody>
</table>

Component 1

Each unique component should be coded separately. A unique component is defined as the set of activates that are dependent on one another. Activities within separate components are not dependent on one another. Select the depth of engagement for the initial description of the component but code all subsequent activities to the main code. Example: A sticker based reminder and a digital reminder could each be implemented independently and would be separate components. Do not add sub-codes

Does not include engagement

Description of interventions that involved no community engagement. Example: The production of cold chain transport of vaccines. Do not add sub-codes

Includes some engagement

Description of activities that involved some community engagement. Example: Health clinics holding vaccination days and posting fliers. Do not add sub-codes

Primary focus is engagement

Description of activities whose primary focus was community engagement. Example: The establishment of community health councils. Do not add sub-codes

Component 10

Does not include engagement

Includes some engagement

Primary focus is engagement

Component 2

Does not include engagement

Includes some engagement

Primary focus is engagement
Component 3
Does not include engagement
Includes some engagement
Primary focus is engagement
Component 4
Does not include engagement
Includes some engagement
Primary focus is engagement
Component 5
Does not include engagement
Includes some engagement
Primary focus is engagement
Component 6
Does not include engagement
Includes some engagement
Primary focus is engagement
Component 7
Does not include engagement
Includes some engagement
Primary focus is engagement
Component 8
Does not include engagement
Includes some engagement
Primary focus is engagement
Component 9
Does not include engagement
Includes some engagement

Primary focus is engagement

<table>
<thead>
<tr>
<th>Category</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Future directions for implementation</td>
<td>Discussion of next steps, how the program could be re-designed, or suggestions for the future. Do not add sub-codes</td>
</tr>
<tr>
<td>Intervention exposure or reach</td>
<td>Discussions of the proportion of the eligible population that the intervention tried to reach. For example, the number of people who received the video (regardless of whether it was actually watched). Do not add sub-codes</td>
</tr>
<tr>
<td>Objective</td>
<td>Statement of the goal, objective, or target of the intervention. Do not add sub-codes</td>
</tr>
<tr>
<td>Participant views of the intervention</td>
<td>Text related to how participants perceived the intervention</td>
</tr>
<tr>
<td>Negative views</td>
<td>Participants did not view the intervention positively</td>
</tr>
<tr>
<td>Positive views</td>
<td>Participants viewed the intervention positively</td>
</tr>
<tr>
<td>Personnel implementing the program</td>
<td>Description of who implemented the program. Do not add sub-codes</td>
</tr>
<tr>
<td>Target group</td>
<td>Description of the target group for this program. Note: This could be different from beneficiary selection if (for example) fathers were provided with information to get them to have their wives immunise their children. Do not add additional sub-codes</td>
</tr>
<tr>
<td>Uptake</td>
<td>Program uptake is defined as initial engagement with the intervention. This could be attending the first meeting, or expressing interest. Use sub-codes when appropriate and only primary code when the text does not fall into the other two. Do not add additional sub-codes</td>
</tr>
<tr>
<td>Assessment of uptake</td>
<td>Description of how uptake was assessed. Example: Attendance sheets were used to collect information on attendance at the first meeting. Do not add sub-codes. Do not add sub-codes</td>
</tr>
<tr>
<td>Description of uptake</td>
<td>Description of uptake. Example: Uptake was high, with 80% of eligible women participating in the intervention. Do not add sub-codes.</td>
</tr>
<tr>
<td>Research design</td>
<td>This set of codes describes the research design and provides information on internal and external validity. Do not add sub-codes.</td>
</tr>
<tr>
<td>Author discussion of external validity</td>
<td>Any explicit discussion by authors. Do not add sub-codes.</td>
</tr>
<tr>
<td>Conflict of interest</td>
<td>Discussion of conflicts of interests. This could include discussions related to the independence of evaluators and data collectors from the implementers and donors. Select the relevant code indicating presence / absence of a conflict. Do not add sub-codes.</td>
</tr>
<tr>
<td>Code</td>
<td>Description</td>
</tr>
<tr>
<td>----------</td>
<td>------------------------------------------------------------------------------</td>
</tr>
<tr>
<td>No</td>
<td>Select any statement directly indicating that there was no conflict of interest. Example: The data collection team was hired as external contractors. Do not add sub-codes.</td>
</tr>
<tr>
<td>Unclear</td>
<td>Select any text that makes a potential conflict of interest unclear. Do not add sub-codes.</td>
</tr>
<tr>
<td>Yes</td>
<td>Select any text that indicates a clear conflict of interest. Example: The implementers were the data collectors. Do not add sub-codes.</td>
</tr>
<tr>
<td>Data collection</td>
<td>Discussion of data collection. Use main code for general discussions and only use sub-codes as appropriate. Do not add sub-codes.</td>
</tr>
<tr>
<td>Response accuracy</td>
<td>Any measure or discussion of response accuracy. Example: We expect that some of our findings may be related to social desirability bias. Do not add sub-codes.</td>
</tr>
<tr>
<td>Retrospective</td>
<td>Text indicating that data collection was retrospective. Include discussion of challenges this may have caused. Do not add additional sub-codes.</td>
</tr>
<tr>
<td>Sampling frame</td>
<td>Descriptions of the sampling frame. Only select text related to how people were ultimately enrolled / selected. Text related to geographic distribution of villages, the method of randomisation (e.g., computer generated vs. paper based), and other details about the preparation for selection is not needed. Example: Women of child bearing age were selected when they presented at the clinic. Do not add sub-codes.</td>
</tr>
<tr>
<td>Data limitations</td>
<td>Discussions of the limitations to the data that pose challenges (e.g., we could not analyse X because of Y). Do not add additional sub-codes.</td>
</tr>
<tr>
<td>Ethical approval</td>
<td>Statement of ethical approval. Do not add sub-codes.</td>
</tr>
<tr>
<td>Future directions for research</td>
<td>Descriptions of future directions for research. Do not add sub-codes.</td>
</tr>
<tr>
<td>Other limitations</td>
<td>Text related to limitations in the research design that are not data related. Example: Failure in randomisation. Example: Failure in randomisation. Do not add sub-codes.</td>
</tr>
<tr>
<td>Uptake and fidelity challenges</td>
<td>This set of codes describes challenges with uptake and fidelity. Most include sub-codes titled &quot;present&quot; and &quot;absent.&quot; In each case, present reflects a statement that this challenge occurred and absent reflects a statement that this challenge was not encountered. Add sub-codes if needed.</td>
</tr>
<tr>
<td>Adherence</td>
<td>Discussion of people (not) completing the intervention activities (e.g., stopped attending training sessions). Could involve discussions of variability in engagement with the intervention or discussions of continued engagement with the intervention. Do not add sub-codes.</td>
</tr>
<tr>
<td>Absent</td>
<td>There were not challenges related to adherence. ie adherence was high.</td>
</tr>
<tr>
<td>Category</td>
<td>Description</td>
</tr>
<tr>
<td>--------------------------------</td>
<td>-----------------------------------------------------------------------------</td>
</tr>
<tr>
<td>Present</td>
<td>There were problems with adherence. ie adherence was low</td>
</tr>
<tr>
<td><strong>Administrative</strong></td>
<td>Challenges related to record keeping, monitoring, and other administrative activities. Do not add sub-codes</td>
</tr>
<tr>
<td>Absent</td>
<td>There were not administrative challenges. Example: All paperwork was conducted quickly and on time</td>
</tr>
<tr>
<td>Present</td>
<td>There were administrative challenges. Ex: Paperwork was finished late and this delayed implementation</td>
</tr>
<tr>
<td><strong>Attrition in the research</strong></td>
<td>People did not complete the research study (eg. did not respond to all rounds of data collection). Do not add sub-codes</td>
</tr>
<tr>
<td>Absent</td>
<td>There was not attrition in the research study. ie attrition was low</td>
</tr>
<tr>
<td>Present</td>
<td>There was significant attrition in the research study. ie attrition was high</td>
</tr>
<tr>
<td><strong>Budget</strong></td>
<td>Challenges due to budget limitations. Do not add sub-codes</td>
</tr>
<tr>
<td>Budget limitations in the intervention</td>
<td>The intervention could not be implemented in a desired way due to budget limitations. Do not add sub-codes</td>
</tr>
<tr>
<td>Absent</td>
<td>There were not budgetary limitations to the intervention. Ex: Due to the generous funding of our donor, we were able to......</td>
</tr>
<tr>
<td>Present</td>
<td>There were budgetary limitations to the intervention. Ex: We had planned X, but due to budget limitations has to Y</td>
</tr>
<tr>
<td>Budget limitations in the research study</td>
<td>The research study could not be implemented as desired due to budget limitations. Do not add sub-codes</td>
</tr>
<tr>
<td>Absent</td>
<td>There were not budgetary limitations to the research study Ex: Due to the generous funding of our donor, we were able to......</td>
</tr>
<tr>
<td>Present</td>
<td>There were budgetary constraints in the research study Ex: We were not able to collect X data due to limited budget</td>
</tr>
<tr>
<td><strong>Contamination</strong></td>
<td>Select the appropriate type of contamination described. Do not add sub-codes</td>
</tr>
<tr>
<td>Contamination between intervention and control</td>
<td>The control group was incidentally exposed to the intervention. Example: Nearby villages were aware of education campaigns conducted.</td>
</tr>
<tr>
<td>Absent</td>
<td>Explicit statements that this contamination did not occur. Do not add sub-codes</td>
</tr>
<tr>
<td>Present</td>
<td>Contamination between intervention and control occurred</td>
</tr>
<tr>
<td><strong>Contamination by other programs</strong></td>
<td>Other ongoing programs may have affected results. Do not add additional sub-codes</td>
</tr>
<tr>
<td>Absent</td>
<td>There was not contamination by other programs Ex: No other similar programs were functioning in the area</td>
</tr>
<tr>
<td>Present</td>
<td>There was contamination by other programs Ex: The Red Cross was conducting a similar program in control villages at the time of our intervention</td>
</tr>
<tr>
<td><strong>Hawthorn effects</strong></td>
<td>Any description of the change in behaviour (real or reported) that is the result of people knowing they are being observed, without any real intent to adopt or maintain these behaviours. Example: Respondents in the UCT group may have been affected by the evaluation process: community awareness about the aims of the project could have affected actual or reported behaviours. Do not add sub-codes</td>
</tr>
<tr>
<td>---</td>
<td>---</td>
</tr>
<tr>
<td><strong>Absent</strong></td>
<td>Hawthorn effects were not observed or expected  Ex: By using an enumeration team that was unaffiliated with the intervention, we expect to have reduced the chances that people adjusted their behaviour simply due to the enumerator's presence</td>
</tr>
<tr>
<td><strong>Present</strong></td>
<td>Hawthorn effects were observed or expected  Ex: Due to significant marketing around the desirability of certain behaviours, people may have changed how they acted in front of enumerators without actually adopting certain behaviours when they were not observed.</td>
</tr>
<tr>
<td><strong>Low implementation fidelity</strong></td>
<td>Deviations in implementation of the intervention from what was planned. Note: this must be from what was planned, not simply what would be considered desirable  Do not add sub-codes</td>
</tr>
<tr>
<td><strong>Absent</strong></td>
<td>Fidelity challenges were not present. The intervention was implemented as intended. Ex: The intervention was implemented as intended</td>
</tr>
<tr>
<td><strong>Present</strong></td>
<td>The intervention was not implemented as intended. Ex: Due to confusion among facilitators, some changed the order in which material was presented.</td>
</tr>
<tr>
<td><strong>Mobilization</strong></td>
<td>Challenges related to mobilizing people to participate  Do not add sub-codes</td>
</tr>
<tr>
<td><strong>Mobilizing in the intervention</strong></td>
<td>People were not willing to engage with the intervention  Do not add sub-codes</td>
</tr>
<tr>
<td><strong>Absent</strong></td>
<td>Low participation was not a challenge observed. ie participation was high.</td>
</tr>
<tr>
<td><strong>Present</strong></td>
<td>The challenge of low participation was encountered. ie participation was low</td>
</tr>
<tr>
<td><strong>Mobilizing in the research</strong></td>
<td>Challenges getting those who participated in the intervention to join the research project. (eg. refuse to be interviewed)  Do not add sub-codes</td>
</tr>
<tr>
<td><strong>Absent</strong></td>
<td>Low participation was not a challenge observed. ie participation was high.</td>
</tr>
<tr>
<td><strong>Present</strong></td>
<td>The challenge of low participation was encountered. ie participation was low</td>
</tr>
<tr>
<td><strong>Other Challenges</strong></td>
<td></td>
</tr>
</tbody>
</table>
## Cost data extraction tool

<table>
<thead>
<tr>
<th>Study Characteristics</th>
<th>Variable Label</th>
<th>Explanation</th>
<th>Source</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Coder</strong></td>
<td>Study ID</td>
<td>This is the study ID - it should match the study ID from the Outcome Mapping Sheet (e.g., SC-SR_1)</td>
<td>Data extraction tool_immunisation</td>
</tr>
<tr>
<td></td>
<td>Study ID Component</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Year</strong></td>
<td>Year</td>
<td>Year published</td>
<td>Data extraction tool_immunisation</td>
</tr>
<tr>
<td><strong>Design</strong></td>
<td>Design</td>
<td>0=Experimental Design (e.g., RCT), 1=Quasi-Experimental Design</td>
<td>Data extraction tool_immunisation</td>
</tr>
<tr>
<td><strong>Country</strong></td>
<td>Country</td>
<td>Country of intervention</td>
<td>Data extraction tool_immunisation</td>
</tr>
<tr>
<td><strong>Study Link</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Which Vaccines were studied?</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>How many vaccine doses were delivered?</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Source of info about vaccines</td>
<td></td>
<td>[Source from CEA Inventory]</td>
<td></td>
</tr>
<tr>
<td><strong>Outcomes</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Implementing partners</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Describe Comparison Group</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Analytical perspective</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Analytical perspective</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Treatment arms</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Comparison</td>
<td>Comparison</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Describe Comparison Group</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Subgroup</td>
<td>Subgroup</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Supplemental material

BMJ Publishing Group Limited (BMJ) disclaims all liability and responsibility arising from any reliance on the information supplied by the author(s). For any queries relating to compliance with ethical requirements, please contact bmjopenqueries@bmj.com.  

<table>
<thead>
<tr>
<th>If yes to subgroup, describe</th>
<th>If yes to subgroup, describe</th>
<th>Free text, describe the subgroup if applicable (e.g., boys, girls). If no subgroup, type N/A</th>
<th>Data extraction tool_immunisation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Source</td>
<td>Source</td>
<td>Note the page number, table number, column, and row you used to extract the data</td>
<td>Data extraction tool_immunisation</td>
</tr>
<tr>
<td>Treatment Effect</td>
<td>Treatment Effect</td>
<td>1=Intention to Treat (ITT), 2=Average Treatment Effect on the Treated (ATET), 3=Average Treatment Effect (ATE) 4 = Local Average Treatment Effect (LATE)</td>
<td>Data extraction tool_immunisation</td>
</tr>
<tr>
<td>Summary of intervention</td>
<td>Summary of intervention</td>
<td>Free text, what is the intervention</td>
<td>Intervention_community engagement_dataset[col E]</td>
</tr>
<tr>
<td>Component #</td>
<td>Component #</td>
<td>List in numbers 1-6</td>
<td>Intervention_community engagement_dataset[col H]</td>
</tr>
<tr>
<td>Description</td>
<td>Description</td>
<td>Description [Col I in Intervention_community engagement_dataset]</td>
<td>Intervention_community engagement_dataset[col I]</td>
</tr>
<tr>
<td>Notes</td>
<td>Notes</td>
<td>Description [Col S in Intervention_community engagement_dataset]</td>
<td>Intervention_community engagement_dataset[col S]</td>
</tr>
<tr>
<td>Exposure to intervention (in months)</td>
<td>Exposure to intervention (in months)</td>
<td>How long is the intervention exposure itself?</td>
<td>Data extraction tool_immunisation</td>
</tr>
<tr>
<td>Evaluation period (in months)</td>
<td>Evaluation period (in months)</td>
<td>The total number of months elapsed between offering an intervention and the point at which an outcome measure is taken post intervention, or as a follow-up measurement. If less than one month, use decimals (e.g., one week would be .25)</td>
<td>Data extraction tool_immunisation</td>
</tr>
<tr>
<td>Post-intervention or change from baseline?</td>
<td>Post-intervention or change from baseline?</td>
<td>0 = Post-intervention, 1 = Change from baseline</td>
<td>Data extraction tool_immunisation</td>
</tr>
<tr>
<td>Source of the Outcome data (only use for outcomes in category JAA)</td>
<td>Source of the Outcome data (only use for outcomes in category JAA)</td>
<td>1=Immunisation Card, 2=Recall, 3=Combination of both immunisation card and recall, 4=Health Admin Data, 5=N/A</td>
<td>Data extraction tool_immunisation</td>
</tr>
<tr>
<td>Author definition of outcome</td>
<td>Author definition of outcome</td>
<td>Free text - How does the author define the outcome?</td>
<td>Data extraction tool_immunisation</td>
</tr>
<tr>
<td>Activities</td>
<td>Activities</td>
<td>What are the main (non-evaluation) activities undertaken by the project? (e.g. program design, targeting, community outreach, training of trainers, training of beneficiaries, M&amp;E, etc.). Describe</td>
<td>Studies</td>
</tr>
<tr>
<td>Activities Data Source</td>
<td>Activities Data Source</td>
<td>Was the source of descriptive information about program activities taken from a table or taken from text? [Text, Table, Other(describe)]</td>
<td>Studies</td>
</tr>
<tr>
<td>Timeline</td>
<td>Timeline</td>
<td>Is an intervention timeline included with the report? [Yes, No, If yes, page #]</td>
<td>Studies</td>
</tr>
<tr>
<td>Ingredients (unit cost of ingredients)</td>
<td>Ingredients (unit cost of ingredients)</td>
<td>Is a list of cost ingredients or inputs for CEA, CBA, or CUA provided? [Yes, No]</td>
<td>Studies</td>
</tr>
<tr>
<td>Question</td>
<td>Response</td>
<td>Studies</td>
<td></td>
</tr>
<tr>
<td>------------------------------------------------------------------------</td>
<td>-----------------------------------------------------------------------------------------------</td>
<td>---------</td>
<td></td>
</tr>
<tr>
<td>Ingredients will be items like &quot;personnel&quot;, &quot;staff&quot;, &quot;travel&quot;, &quot;office support&quot;, &quot;program materials&quot;, etc. Note that searches for the unit costs of activities should include text searches as well as searches of tables.</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Are unit costs reported?</td>
<td>Look for a table with units or unit costs reported. Unit costs give the value of individual items that were used during the intervention, for example the unit cost of a staff person, e.g. a &quot;health specialist&quot; would be given in terms of monthly wages, or compensation or salary. Note that searches for the unit costs of activities should include text searches as well as searches of tables. [Yes, No, If Yes Page #]</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cost data sources</td>
<td>Was a source of cost data reported? (i.e. financial reports or accounting systems of implementing NGOs or donors (expenditures), NGO or donor (budgets), (market prices) Describe (or copy from text). Clarify - expenditure is what's actually spent, budget is what's planned for spending / not audited, often incorrect and not updated.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Costing methodology</td>
<td>Was a method of costing described? (i.e. key works: gross costing, micro-costing, ingredients method, activity-based costing method) [if yes, what was the method].</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Type of efficiency analysis (see glossary)</td>
<td>Indicate (CBA/CEA/CUA/cost-minimization, or innovative approaches: Social Return on Investment (SROI), Multi-Criteria Appraisal (MCA). Response options [CBA, CEA, CUA, SROI, MCA, or N/A if only &quot;total cost&quot;, also ok to insert descriptive results that do not conform, e.g. “cost transfer ratio”]</td>
<td>Studies</td>
<td></td>
</tr>
<tr>
<td>Any cost-effectiveness analysis?</td>
<td>Source [from CEA Inventory]</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Any quant discussion of costs? (Full program)</td>
<td>Source [from CEA Inventory]</td>
<td></td>
<td></td>
</tr>
<tr>
<td>If yes, what type? [Page number]</td>
<td>Source [from CEA Inventory]</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Currency of cost reporting</td>
<td>What is the currency in which costs were reported, i.e. USD, Rwandan Francs, etc.</td>
<td>Studies</td>
<td></td>
</tr>
<tr>
<td>Exchange rate</td>
<td>What exchange rate was used for currency adjustments? [e.g. 1.2 USD to 1 Euro]</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Exchange rate year (estimated)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Exchange rate date</td>
<td>What was the date, month, year of the exchange rate used? [If the exchange rate was calculated as an average over the implementation period, copy verbatim from text the definition of average exchange rate used]</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Question</td>
<td>Answer Format</td>
<td>Studies</td>
<td></td>
</tr>
<tr>
<td>-------------------------------------------------------------------------</td>
<td>---------------</td>
<td>---------</td>
<td></td>
</tr>
<tr>
<td>Was there discussion of a discount rate?</td>
<td>[Yes, No]</td>
<td>Studies</td>
<td></td>
</tr>
<tr>
<td>Discount rate</td>
<td></td>
<td>Studies</td>
<td></td>
</tr>
<tr>
<td>If no discount rate is applied, was any justification provided?</td>
<td>Paste explanation provided in text. We consider any acknowledgement of the discount rate to be justification. For example, ok to report “we did not discount costs or benefits” as justification</td>
<td>Studies</td>
<td></td>
</tr>
<tr>
<td>Inflation adjustment</td>
<td></td>
<td>Studies</td>
<td></td>
</tr>
<tr>
<td>Inflation adjustment of cost reported</td>
<td></td>
<td>Studies</td>
<td></td>
</tr>
<tr>
<td>Base year of the costing</td>
<td>base_year</td>
<td>Studies</td>
<td></td>
</tr>
<tr>
<td>Average cost per number of participants</td>
<td></td>
<td>Studies</td>
<td></td>
</tr>
<tr>
<td>What is the denominator used in average intervention cost?</td>
<td>ac_denominator</td>
<td>Studies</td>
<td></td>
</tr>
<tr>
<td>Average cost per DALY or QALY?</td>
<td></td>
<td>Studies</td>
<td></td>
</tr>
<tr>
<td>Total intervention cost (excluding vaccine cost)</td>
<td>total_cost</td>
<td>Studies</td>
<td></td>
</tr>
<tr>
<td>Did the total intervention cost reported above include the cost of vaccines?</td>
<td>vac_incl</td>
<td>Studies</td>
<td></td>
</tr>
<tr>
<td>Total cost description</td>
<td>total_cost_desc</td>
<td>Studies</td>
<td></td>
</tr>
<tr>
<td>Government contribution included in total cost?</td>
<td>tc_gov_contb</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Is more than one total cost given in the report?</td>
<td>[Yes, No, If Yes, please describe additional &quot;total cost&quot; that is reported]</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Number of immunisations</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cost per vaccine delivered, cost per additional immunisation administered or cost per dose (of vaccine). Type of vaccine is given in col I</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Description and page number of cost per additional vaccine delivered.</td>
<td>Note if average costs per additional vaccine or marginal costs per additional vaccine is reported, where marginal costs should exclude fixed costs.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>---</td>
<td>---</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cost per additional child immunised_ All</td>
<td>cpc_all</td>
<td>[Yes, No, If Yes, please report cost per additional child immunised_all]</td>
<td></td>
</tr>
<tr>
<td>Description and page number of cost per additional CHILD Immunised</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cost per additional child immunised_DPT3</td>
<td>cpc_dpt3</td>
<td>[Yes, No, If Yes, please report cost per additional child immunised with DPT3]. Please make a note if only DPT1 was reported.</td>
<td></td>
</tr>
<tr>
<td>Cost per additional child immunised_Measles</td>
<td>cpc_measles</td>
<td>[Yes, No, If Yes, please report cost per additional child immunised with measles]</td>
<td></td>
</tr>
<tr>
<td>Baseline immunisation coverage of the target population_All</td>
<td>What was the reported baseline rate of immunisation (all immunisations) in the target population? [Percentage, or range]</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Was baseline immunisation coverage of All immunisations reported for the target population, general population, or sample?</td>
<td>Was baseline immunisation coverage of All immunisations reported for the target population, general population, or sample? [target population, general population, sample]</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Baseline immunisation coverage of the target population_DPT3</td>
<td>What was the reported baseline rate of immunisation (DPT3) in the target population? [Percentage, or range]</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Was baseline immunisation coverage of DPT3 reported for the target population, general population, or sample?</td>
<td>Was baseline immunisation coverage of DPT3 reported for the target population, general population, or sample? [target population, general population, sample]</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Baseline immunisation coverage of the target population_Measles</td>
<td>What was the reported baseline rate of immunisation (Measles) in the target population? [Percentage, or range]</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Final immunisation coverage of the target population_All</td>
<td>What was the reported final rate of immunisation (all immunisations) in the target population? [Percentage, or range]</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Final immunisation coverage of the target population_DPT3</td>
<td>What was the reported final rate of immunisation (DPT3) in the target population? [Percentage, or range]</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Final immunisation coverage of the target population_Measles</td>
<td>What was the reported final rate of immunisation (Measles) in the target population? [Percentage, or range]</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Non-compliance</td>
<td>Was there any evidence of non-compliance with treatment assignment. Non-compliance is where individuals assigned to treatment do not take the treatment, or the case where individuals assigned to control DO take the treatment. Ok to keyword search on &quot;non-compliance&quot;, &quot;non compliance&quot;, &quot;noncompliance&quot; [Yes, No] [If Yes, copy related text on non-compliance]</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Studies
<table>
<thead>
<tr>
<th>Spillover</th>
<th>Was there any evidence of spillover where individuals assigned to control receive a benefit or incur a ‘cost’ that results from the treatment intervention? [Yes, No] [If Yes, copy related text on spillovers detected]</th>
<th>Studies</th>
</tr>
</thead>
<tbody>
<tr>
<td>Attrition or drop out</td>
<td>Was there evidence that participants dropped out of the study before follow-up could be completed? [Yes, No] [If Yes, copy related text on drop outs]</td>
<td>Studies</td>
</tr>
<tr>
<td>Vaccination coverage_All</td>
<td>JAA01. Full routine immunisation for children</td>
<td>Binary measure of whether or not children have received all routine vaccinations for the relevant country or region.</td>
</tr>
<tr>
<td>Vaccination coverage_DPT3</td>
<td>JAA05. DPT3</td>
<td>Binary measure of whether or not children have received the third dose of the DPT or pentavalent vaccine. If the study does not specifically say “DPT3” (or “pentavalent 3”), but refers to “complete DPT/penta vaccination” or something similar, then use this code.</td>
</tr>
<tr>
<td>Vaccination coverage_Measles</td>
<td>JAA11. Measles</td>
<td>Binary measure of whether or not children have received the measles vaccine.</td>
</tr>
<tr>
<td>Costs for significant outcomes only?</td>
<td>Does the report mention that costs will be reported only where significant outcomes are observed? [Yes, No]</td>
<td></td>
</tr>
</tbody>
</table>

### List of coded variables

<table>
<thead>
<tr>
<th>VARIABLE LABEL</th>
<th>EXPLANATION</th>
</tr>
</thead>
<tbody>
<tr>
<td>Study ID</td>
<td>This is the study ID - it should match the study ID from the Outcome Mapping Sheet (e.g., SC-SR_1)</td>
</tr>
<tr>
<td>Estimate ID</td>
<td>The estimate ID will provide a specific number for each effect size extracted and should include the original study number, underscore, then the unique ID number (e.g., SC-SR1_1, SC-SR1_2 and so on)</td>
</tr>
<tr>
<td>Author</td>
<td>For 1 author: leading author last name (e.g., Gomez) For 2 authors: both author last names with ampersand in between (e.g., Smith &amp; Bahn) For 3 or more authors: leading author last name followed by et al. (e.g., Gupta et al.)</td>
</tr>
<tr>
<td>Year</td>
<td>Year published</td>
</tr>
<tr>
<td>Design</td>
<td>0=Experimental Design (e.g., RCT), 1=Quasi-Experimental Design</td>
</tr>
<tr>
<td>How Counterfactual is Chosen</td>
<td>Free text (e.g., random control trial, propensity score matching, etc) - Multiple codes are ok</td>
</tr>
<tr>
<td>Analysis type for this effect size</td>
<td>Free text, what type of analysis was used (Regression, 2SLS, ANCOVA, etc.)- Multiple codes are ok</td>
</tr>
<tr>
<td>Country</td>
<td>Country of intervention</td>
</tr>
<tr>
<td>Estimate Type</td>
<td>Type of data for this effect size: 1 = Continuous - means and SDs, 2 = Continuous - mean difference and SD, 3 = Dichotomous outcome - proportions, 4 = Regression data - dichotomous outcome (e.g., logistic regression)</td>
</tr>
<tr>
<td><strong>Comparison</strong></td>
<td>5 = Regression data - continuous outcome (e.g., linear regression)</td>
</tr>
<tr>
<td>---------------</td>
<td>---------------------------------------------------------------</td>
</tr>
<tr>
<td><strong>Describe Comparison Group</strong></td>
<td>If answer above is (1) no intervention, type N/A, if (2) Other Intervention, list what intervention the control group is receiving, if (3) Pipeline control, report when the control group will receive the intervention in relation to the treatment group (e.g., one year later)</td>
</tr>
<tr>
<td><strong>Subgroup</strong></td>
<td>Is this analysis of a subgroup? 0=no, 1=yes</td>
</tr>
<tr>
<td><strong>If yes to subgroup, describe</strong></td>
<td>Free text, describe the subgroup if applicable (e.g., boys, girls). If no subgroup, type N/A</td>
</tr>
<tr>
<td><strong>Source</strong></td>
<td>Note the page number, table number, column, and row you used to extract the data</td>
</tr>
<tr>
<td><strong>Treatment Effect</strong></td>
<td>1=Intention to Treat (ITT), 2=Average Treatment Effect on the Treated (ATET), 3=Average Treatment Effect (ATE) 4 = Local Average Treatment Effect (LATE)</td>
</tr>
<tr>
<td><strong>Intervention</strong></td>
<td>Free text, what is the intervention</td>
</tr>
<tr>
<td><strong>Exposure to intervention (in months)</strong></td>
<td>The total number of months elapsed between offering an intervention and the point at which an outcome measure is taken post intervention, or as a follow-up measurement. If less than one month, use decimals (e.g., one week would be .25)</td>
</tr>
<tr>
<td><strong>Evaluation period (in months)</strong></td>
<td>0 = Post-intervention, 1 = Change from baseline</td>
</tr>
<tr>
<td><strong>Source of the Outcome data (only use for outcomes in category JAA)</strong></td>
<td>1=Immunization Card, 2=Recall, 3=Combination of both immunization card and recall, 4=Health Admin Data, 5=N/A, 6 = unclear</td>
</tr>
<tr>
<td><strong>Author definition of outcome</strong></td>
<td>Free text - How does the author define the outcome?</td>
</tr>
<tr>
<td><strong>Intervention</strong></td>
<td>Coded as engagement as intervention, engagement in the intervention design, engagement in implementation autonomy, or multiple engagement types (place a 1 in the appropriate column)</td>
</tr>
<tr>
<td><strong>If yes to multiple engagement types</strong></td>
<td>Identify which engagement types are used.</td>
</tr>
<tr>
<td><strong>Implemented by the government</strong></td>
<td>0 = no, 1 = yes</td>
</tr>
<tr>
<td><strong>Were new cadres of health workers formed?</strong></td>
<td>0 = no, 1 = yes</td>
</tr>
<tr>
<td><strong>Was vaccination hesitency identified as a barrier?</strong></td>
<td>0 = no, 1 = yes</td>
</tr>
<tr>
<td><strong>Baseline rates of full immunisation</strong></td>
<td>record percentage vaccinated</td>
</tr>
<tr>
<td><strong>Baseline rates of DPT3</strong></td>
<td>record percentage vaccinated</td>
</tr>
</tbody>
</table>

**OUTCOME CODES**
Code 1 under any applicable columns.

**EFFECT SIZE DATA EXTRACTION**
Record 0='no' if an increase is good, record 1='yes' if a decrease is good and the sign needs to be reversed.

**Unit of analysis**
What is the unit of analysis? UOA for this effect size: 1= Individual, 2= Household, 3= Group (e.g.,...
<table>
<thead>
<tr>
<th>Variable</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>mean_t</td>
<td>Outcome mean for the treatment group</td>
</tr>
<tr>
<td>sd_t</td>
<td>Outcome standard deviation for treatment group</td>
</tr>
<tr>
<td>mean_c</td>
<td>Outcome mean for the comparison group</td>
</tr>
<tr>
<td>sd_c</td>
<td>Outcome standard deviation for control group</td>
</tr>
<tr>
<td>mean_overall_diff</td>
<td>Overall mean difference (treatment - control)</td>
</tr>
<tr>
<td>diff se</td>
<td>Standard error of the overall mean difference</td>
</tr>
<tr>
<td>Diff_t</td>
<td>t-statistic of mean difference</td>
</tr>
<tr>
<td>Odds</td>
<td>Odds ratio reported in the study</td>
</tr>
<tr>
<td>OR_se</td>
<td>Odds ratio standard error reported in the study</td>
</tr>
<tr>
<td>Risk ratio</td>
<td>Risk ratio reported in study</td>
</tr>
<tr>
<td>RR_se</td>
<td>Risk ratio standard error</td>
</tr>
<tr>
<td>reg_coeff</td>
<td>Report the regression coefficient of the treatment effect</td>
</tr>
<tr>
<td>reg_SE</td>
<td>Report the associated standard error of the regression coefficient</td>
</tr>
<tr>
<td>reg_t</td>
<td>Report the associated t statistic of the effect size (coefficient/SE)</td>
</tr>
<tr>
<td>Exact p value</td>
<td>Exact p value if given, if not, record as written in the manuscript (e.g., p &lt; .001, or p &gt; .05)</td>
</tr>
<tr>
<td>clust_t</td>
<td>Number of clusters - treatment group</td>
</tr>
<tr>
<td>clust_c</td>
<td>Number of clusters - control group</td>
</tr>
<tr>
<td>clust_T</td>
<td>Number of clusters - total sample</td>
</tr>
<tr>
<td>n_t</td>
<td>Sample size - treatment group</td>
</tr>
<tr>
<td>n_c</td>
<td>Sample size - control group</td>
</tr>
<tr>
<td>n_T</td>
<td>Sample size - total sample</td>
</tr>
<tr>
<td>periods (1 if cross sectional)</td>
<td>Record how many periods of evaluation there are (e.g., cross section is 1, panel data with 3 measurements is 3)</td>
</tr>
<tr>
<td>Treatment Variable</td>
<td>Record the treatment variable as written in the model (e.g., the variable name the author uses, such as &quot;Intervention x Time&quot;)</td>
</tr>
<tr>
<td>dataset</td>
<td>Record if data comes from an identified dataset</td>
</tr>
<tr>
<td>coder</td>
<td>Record your name</td>
</tr>
<tr>
<td>Notes</td>
<td>Record any notes important for the team</td>
</tr>
<tr>
<td>ROB Category</td>
<td>0 = low risk of bias, 1 = some concerns, 2 = high risk of bias</td>
</tr>
</tbody>
</table>
### Appendix 7: Risk of bias assessment tools

#### Quantitative risk of bias tool

<table>
<thead>
<tr>
<th>Code</th>
<th>Question</th>
<th>Coding format</th>
<th>Criteria</th>
</tr>
</thead>
<tbody>
<tr>
<td>General</td>
<td>General ID</td>
<td>EPPI ID</td>
<td></td>
</tr>
<tr>
<td>Genera</td>
<td>Study first author</td>
<td>Open answer</td>
<td></td>
</tr>
<tr>
<td>General</td>
<td>Time taken to complete assessment</td>
<td>Minutes</td>
<td></td>
</tr>
<tr>
<td>General</td>
<td>Design type: What type of study design is used?</td>
<td>1 = Randomised controlled trial (RCT) (random assignment to households/individuals) or quasi-RCT 2 = Cluster-RCT (quasi-RCT) 3 = Pseudo-RCT</td>
<td>-</td>
</tr>
<tr>
<td>General</td>
<td>Methods used for analysis: Which methods are used to control for selection bias and confounding?</td>
<td>1 = Statistical matching (PSM, CEM, covariate matching) 2 = Difference in differences (DID) estimation methods 3 = IV-regression (2-stage least squares or bivariate probit) 4 = Heckman selection model 5 = Fixed effects or random effects regression 6 = Covariate adjusted estimation 7 = Propensity weighted regression 8 = Comparison of means 9 = Other (please state)</td>
<td>-</td>
</tr>
<tr>
<td>General</td>
<td>Design and analysis method description</td>
<td>Open answer</td>
<td>Briefly describe the study design and analysis method undertaken by the authors.</td>
</tr>
<tr>
<td>General</td>
<td>Study population</td>
<td>Open answer</td>
<td>Provide any details in the paper that describe how the study population was selected, answering the question: what is the sampling strategy to recruit participants from that population into the evaluation?</td>
</tr>
<tr>
<td>General</td>
<td>Type of comparison group</td>
<td>1 = No intervention (service delivery as usual) 2 = Other intervention 3 = Pipeline (wait-list) control (still service delivery as usual)</td>
<td>Indicate type of comparison group</td>
</tr>
<tr>
<td>General</td>
<td>Type of comparison group (if other)</td>
<td>Open answer</td>
<td></td>
</tr>
<tr>
<td>General</td>
<td>Ethical clearance</td>
<td>Open answer</td>
<td>Provide any details of ethical research clearances granted. Report unclear if this information is not available.</td>
</tr>
<tr>
<td>---------</td>
<td>------------------</td>
<td>-------------</td>
<td>----------------------------------------------------------------------------------------------------------------</td>
</tr>
<tr>
<td>General</td>
<td>Study registration</td>
<td>Open answer</td>
<td>Provide any details of study registration, including registry IDs, etc.</td>
</tr>
<tr>
<td>1: Bias arising from randomisation process</td>
<td>1.1 Was the allocation sequence random?</td>
<td>(1) Yes; (2) Probably yes; (3) Probably no; (4) No; (5) No information;</td>
<td>1.1 Was the allocation sequence random? Answer ‘Yes’ if a random component was used in the sequence generation process. Examples include computer-generated random numbers; reference to a random number table; coin tossing; shuffling cards or envelopes; throwing dice; or drawing lots. Minimization is generally implemented with a random element (at least when the scores are equal), so an allocation sequence that is generated using minimization should generally be considered to be random. Answer ‘No’ if no random element was used in generating the allocation sequence or the sequence is predictable. Examples include alternation; methods based on dates (of birth or admission); patient record numbers; allocation decisions made by clinicians or participants; allocation based on the availability of the intervention; or any other systematic or haphazard method. Answer ‘No information’ if the only information about randomization methods is a statement that the study is randomized. Note: In some situations a judgement may be made to answer ‘Probably no’ or ‘Probably yes’. For example, if the study was large, conducted by an independent trials unit or carried out for regulatory purposes, it may be reasonable to assume that the sequence was random. Alternatively, if other (contemporary) trials by the same investigator team have clearly used non-random sequences, it might be reasonable to assume that the current study was done using similar methods.</td>
</tr>
<tr>
<td>1: Bias arising from randomisation process</td>
<td>Question 1.1 answer justification</td>
<td>Open answer</td>
<td>Justification for coding decision (Include a brief summary of justification for rating, mentioning your response to all sub questions, cite relevant pages).</td>
</tr>
<tr>
<td>1: Bias arising from randomisation process</td>
<td>1.2 Did baseline differences between intervention groups suggest a problem with the randomisation process?</td>
<td>(1) Yes; (2) Probably yes; (3) Probably no; (4) No; (5) No information;</td>
<td>1.2 Did baseline differences between intervention groups suggest a problem with the randomization process? Note that differences that are compatible with chance do not lead to a risk of bias.</td>
</tr>
</tbody>
</table>
Answer ‘No’ if no imbalances are apparent or if any observed imbalances are compatible with chance.

Answer ‘Yes’ if there are imbalances that indicate problems with the randomization process, including:
(1) substantial differences between intervention group sizes, compared with the intended allocation ratio; or
(2) a substantial excess in statistically significant differences in baseline characteristics between intervention groups, beyond that expected by chance (i.e., on more than 10% of the characteristics reported); or
(3) imbalance in one or more key prognostic factors, or baseline measures of outcome variables, that is very unlikely to be due to chance and for which the between-group difference is big enough to result in bias in the intervention effect estimate.

Also answer ‘Yes’ if there are other reasons to suspect that the randomization process was problematic, such as:
(4) excessive similarity in baseline characteristics that is not compatible with chance; or
(5) surprising absence of one or more key baseline characteristics that would be expected to be reported.

Answer ‘No information’ when there is no useful baseline information available (e.g. abstracts, or studies that reported only baseline characteristics of participants in the final analysis).

Note: Trialists may undertake analyses that attempt to deal with flawed randomization by controlling for imbalances in prognostic factors at baseline. To remove the risk of bias caused by problems in the randomization process, it would be necessary to know, and measure, all the prognostic factors that were imbalanced at baseline. It is unlikely that all important prognostic factors are known and measured, so such analyses will at best reduce the risk of bias.

<table>
<thead>
<tr>
<th>1: Bias arising from randomisation process</th>
<th>Question 1.2 answer justification</th>
<th>Open answer</th>
<th>Justification for coding decision</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td>(Include a brief summary of justification for rating, mentioning your response to all sub questions, cite relevant pages).</td>
</tr>
</tbody>
</table>

| 1: Bias arising from allocation or identification | Assignment mechanism: Was the allocation or identification | (0) Low risk of bias | Utilize your answers for questions 1.1 and 1.2 to determine the overall score. |

|                                  |                                    | (1) Medium risk of bias/Some concerns |                                    |

### randomisation process

<table>
<thead>
<tr>
<th>(2) High risk of bias</th>
<th>Justification for coding decision</th>
</tr>
</thead>
<tbody>
<tr>
<td>Open answer</td>
<td>(Include a brief summary of justification for rating, mentioning your response to all sub questions, cite relevant pages).</td>
</tr>
</tbody>
</table>

#### 1: Assignment mechanism - Justification

<table>
<thead>
<tr>
<th>Assignment justification</th>
<th>1=Yes</th>
<th>2=No</th>
<th>3=Not reported/unclear</th>
<th>4=Not applicable</th>
</tr>
</thead>
<tbody>
<tr>
<td>Open answer</td>
<td>Score &quot;Yes&quot; if Unit of Analysis = Unit of Randomization OR if UoA ≠ UoR and standard errors are clustered at the UoR level OR data is collapsed to the UoR level. Score &quot;Not reported/unclear&quot; if not enough information is provided on the way the standard errors were calculated or what the unit of analysis is. Score &quot;Not applicable&quot; if it is not a cluster RCT. Score &quot;No&quot; otherwise.</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

#### 2: Unit of analysis - Assessment

<table>
<thead>
<tr>
<th>Unit of analysis: Is unit of analysis in cluster allocation addressed in standard error calculation?</th>
<th>1=Yes</th>
<th>2=No</th>
<th>3=Not reported/unclear</th>
<th>4=Not applicable</th>
</tr>
</thead>
<tbody>
<tr>
<td>Open answer</td>
<td>Score &quot;Yes&quot; if Unit of Analysis = Unit of Randomization OR if UoA ≠ UoR and standard errors are clustered at the UoR level OR data is collapsed to the UoR level. Score &quot;Not reported/unclear&quot; if not enough information is provided on the way the standard errors were calculated or what the unit of analysis is. Score &quot;Not applicable&quot; if it is not a cluster RCT. Score &quot;No&quot; otherwise.</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

#### 2: Unit of analysis - Justification

<table>
<thead>
<tr>
<th>Question 2 answer justification (provide page numbers)</th>
<th>Open answer</th>
</tr>
</thead>
<tbody>
<tr>
<td>-</td>
<td>-</td>
</tr>
</tbody>
</table>

#### 3. Bias due to missing outcome data (panel data)

<table>
<thead>
<tr>
<th>3.1. Was there attrition, if yes, then what was the rate in treatment and control/comparison groups?</th>
<th>3.1. Was there attrition, if yes, then what was the rate in treatment and control/comparison groups?</th>
</tr>
</thead>
<tbody>
<tr>
<td>(1) Yes;</td>
<td>In case of attrition, answer 'yes' or 'probably yes', otherwise answer 'no' or 'probably no'.</td>
</tr>
<tr>
<td>(2) Probably yes;</td>
<td>If there is no discussion on attrition then score 'no information'.</td>
</tr>
<tr>
<td>(3) Probably no;</td>
<td>Specify the rate of attrition by the study arms.</td>
</tr>
<tr>
<td>(4) No;</td>
<td>Specify the attrition rate assumed for power calculations.</td>
</tr>
<tr>
<td>(5) No information;</td>
<td></td>
</tr>
</tbody>
</table>

---

<table>
<thead>
<tr>
<th>3. Bias due to missing outcome data – Justification (panel data)</th>
<th>Question 3.1 answer justification</th>
<th>Open answer</th>
<th>Has the study discussed the ability of the follow-up study to detect the hypothesised outcome effect with the sample size attained?</th>
</tr>
</thead>
<tbody>
<tr>
<td>3. Bias due to missing outcome data (panel data)</td>
<td>3.2. Does the study establish that attrition is randomly distributed (e.g. by examining correlation with determinants of outcomes, in both treatment and comparison groups or by presenting data showing balance on key characteristics across treatment and control)?</td>
<td>(1) Yes; (2) Probably yes; (3) Probably no; (4) No; (5) No information;</td>
<td>Justification for coding decision (Include a brief summary of justification for rating, mentioning your response to all sub questions, cite relevant pages).</td>
</tr>
<tr>
<td>3. Bias due to missing outcome data (panel data)</td>
<td>3.2. Does the study establish that attrition is randomly distributed (e.g. by examining correlation with determinants of outcomes, in both treatment and comparison groups or by presenting data showing balance on key characteristics across treatment and control)?</td>
<td>Score 'yes' or 'probably yes' if balance on key characteristics between attritors and non-attritors for each intervention group (including control) using baseline data has been presented and finds that there is no systematic difference on at least observable dimensions. Score 'no' or 'probably no' if the study finds that there is a systematic difference on observables between attritors and non-attritors in any of the intervention groups. If the baseline balance table between attritors and non-attritors is not presented or discussed then score 'no information'.</td>
<td></td>
</tr>
<tr>
<td>3. Bias due to missing outcome data – Justification (panel data)</td>
<td>3.3. If there is non-random attrition, have authors used convincing statistical techniques to identify and adjust for the attrition bias?</td>
<td>(1) Yes; (2) Probably yes; (3) Probably no; (4) No; (5) No information; (6) Not applicable</td>
<td>Justification for coding decision (Include a brief summary of justification for rating, mentioning your response to all sub questions, cite relevant pages).</td>
</tr>
<tr>
<td>3. Bias due to missing outcome data – (panel data)</td>
<td>3.3. If there is non-random attrition, have authors used convincing statistical techniques to identify and adjust for the attrition bias?</td>
<td></td>
<td>If the study has used convincing parametric or non-parametric techniques (like ignorable maximum likelihood, Manski-Lee bounds) to identify and adjust for attrition bias, then score 'yes' or 'probably yes'. If the techniques used are not convincing then score 'no' or 'probably no'. If the study does not discuss using any statistical techniques for identifying and adjusting for attrition bias then score 'no information'.</td>
</tr>
</tbody>
</table>
If there is no non-random attrition, then this question does not apply and the score should be 'not applicable'.

<table>
<thead>
<tr>
<th>Question 3.3 answer justification</th>
<th>Open answer</th>
<th>Justification for coding decision (Include a brief summary of justification for rating, mentioning your response to all sub questions, cite relevant pages).</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bias due to missing outcome data – Justification (panel data)</td>
<td>Open answer</td>
<td>Justification for coding decision (Include a brief summary of justification for rating, mentioning your response to all sub questions, cite relevant pages).</td>
</tr>
<tr>
<td>(0) Low risk of bias (1) Medium risk of bias/Some concerns (2) High risk of bias</td>
<td>Utilize your answers for questions 3.1 to 3.3 to determine the overall score.</td>
<td></td>
</tr>
<tr>
<td>3. Bias due to missing outcome data (repeated cross section)</td>
<td>(1) Yes; (2) Probably yes; (3) Probably no; (4) No; (5) No information;</td>
<td>3.1. Was there difference in sample size across baseline and follow-up surveys among treatment and control groups? If yes, what was the rate of difference among treatment and control/comparison groups? Was it lower or higher than baseline? In case of difference, answer 'yes' or 'probably yes', otherwise answer 'no' or 'probably no'. If there is no information to assess the differences, then score 'no information'. Specify the rate of difference by the study arms. Specify the non-compliance and/or attrition rate assumed for power calculations. Has the study discussed the ability of the follow-up study to detect the hypothesised outcome effect with the sample size attained?</td>
</tr>
<tr>
<td>Question 3.1 answer justification</td>
<td>Open answer</td>
<td>Justification for coding decision (Include a brief summary of justification for rating, mentioning your response to all sub questions, cite relevant pages).</td>
</tr>
<tr>
<td>3. Bias due to missing outcome data – Justification (repeated cross section)</td>
<td>Open answer</td>
<td>Justification for coding decision (Include a brief summary of justification for rating, mentioning your response to all sub questions, cite relevant pages).</td>
</tr>
<tr>
<td>3.2. Does the study establish that the sampling frame for baseline and follow-up surveys is same? Are the clusters (eg health center) and sub-clusters (eg. villages) the same across baseline and follow-up surveys?</td>
<td>Open answer</td>
<td>Justification for coding decision (Include a brief summary of justification for rating, mentioning your response to all sub questions, cite relevant pages).</td>
</tr>
</tbody>
</table>
| data (repeated cross section) | surveys is same? Are the clusters (eg health center) and sub-clusters (eg. villages) the same across baseline and follow-up surveys? | (4) No; (5) No information; | Score 'yes' or 'probably yes' if sampling frame (clusters and sub-clusters) are the same for baseline and follow-up surveys
Score 'no' or 'probably no' if the clusters or sub-clusters were different across baseline and follow-up surveys.
If there is no information on sampling frame across baseline or follow-up surveys, then score 'no information'. |
| 3. Bias due to missing outcome data – Justification (repeated cross section) | Question 3.2 answer justification | Open answer | Justification for coding decision
(Include a brief summary of justification for rating, mentioning your response to all sub questions, cite relevant pages). |
| 3. Bias due to missing outcome data –(repeated cross section) | 3.3. Was the sample selection of respondents carried out in the same way as in the baseline? Was the intended number of respondents in the clusters/sub-clusters same across baseline and follow-up surveys? | (1) Yes; (2) Probably yes; (3) Probably no; (4) No; (5) No information; | 3.3. Was the sample selection of respondents carried out in the same way as in the baseline? Was the intended number of respondents in the clusters/sub-clusters same across baseline and follow-up surveys?
Score 'yes' or 'probably yes' if the study used same listing process for selection of respondents and their intended number was same across baseline and follow-up surveys,
Score 'no' or 'probably no' if the study used different listing process for selection of respondents or their intended number was different across baseline and follow-up surveys.
If there is no information on the selection process of respondents then score 'no information'. |
| 3. Bias due to missing outcome data - Justification (repeated cross section) | Question 3.3 answer justification | Open answer | Justification for coding decision
(Include a brief summary of justification for rating, mentioning your response to all sub questions, cite relevant pages). |
| 3. Bias due to missing outcome data – Assessment | Bias due to missing outcome data | (0) Low risk of bias (1) Medium risk of bias/Some concerns (2) High risk of bias | Utilize your answers for questions 3.1 to 3.3 to determine the overall score using the criteria below |

---

<table>
<thead>
<tr>
<th>Bias due to missing outcome data – Justification (repeated cross section)</th>
<th>Bias due to missing outcome data justification</th>
<th>Open answer</th>
<th>Justification for coding decision (Include a brief summary of justification for rating, mentioning your response to all sub questions, cite relevant pages).</th>
</tr>
</thead>
<tbody>
<tr>
<td>3. Bias due to missing outcome data - Assessment</td>
<td>(0) Low risk of bias</td>
<td>Justification for coding decision</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(1) Medium risk of bias/Some concerns</td>
<td>Utilize your answers for questions 3.1 to 3.3 to determine the overall score.</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(2) High risk of bias</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>3. Bias due to missing outcome data - Justification</th>
<th>Bias due to missing outcome data justification</th>
<th>Open answer</th>
<th>Justification for coding decision (Include a brief summary of justification for rating, mentioning your response to all sub questions, cite relevant pages).</th>
</tr>
</thead>
<tbody>
<tr>
<td>4: Bias due to deviations from intended interventions - Assessment</td>
<td>4.1. Was the study adequately protected against spill-overs/contamination?</td>
<td>(1) Yes;</td>
<td>4.1. Was the study adequately protected against spill-overs/contamination?</td>
</tr>
<tr>
<td></td>
<td>(2) Probably yes;</td>
<td>(3) Probably no;</td>
<td>Score “Yes” if the intervention is unlikely to spill-over to comparisons (e.g. participants and non-participants are geographically and/or socially separated from one another and general equilibrium effects are not likely) and that the treatment and comparisons are isolated from other interventions which might explain changes in outcomes.</td>
</tr>
<tr>
<td></td>
<td>(4) No;</td>
<td>(5) No information;</td>
<td>Score “No” if allocation was at the individual or classroom level and there are likely spill-overs within households and communities which are not controlled for (such as using trasferable vouchers), or other interventions likely to affect outcomes operating at the same time in either group.</td>
</tr>
<tr>
<td>4: Bias due to deviations from intended interventions - Justification</td>
<td>Question 4.1 answer justification</td>
<td>Open answer</td>
<td>Justification for coding decision (Include a brief summary of justification for rating, mentioning your response to all sub questions, cite relevant pages).</td>
</tr>
<tr>
<td>Question</td>
<td>Yes</td>
<td>Probably yes</td>
<td>Probably no</td>
</tr>
<tr>
<td>----------</td>
<td>-----</td>
<td>--------------</td>
<td>-------------</td>
</tr>
<tr>
<td>4.2. Was the process of being monitored free from bias?</td>
<td>Score &quot;Yes&quot; if the authors state explicitly that the process of monitoring the intervention is blinded, or argue convincingly why it is not likely that intervention delivery could affect the performance of participants in treatment and comparison groups in different ways (such as resulting in Hawthorne or John Henry effects). or there is nothing in the surveys that might have given the control participants an idea of what the other group might receive (e.g. provide cash transfer to the intervention area participants or reveal information that they did not have before or they did but there is no risk that this has changed their behaviours;</td>
<td>Score &quot;No&quot; if the authors do not use an appropriate method to prevent Hawthorne and John Henry Effects (e.g. blinding of monitoring or other methods to ensure consistent monitoring across groups) and there is a risk that the intervention delivery or survey process could have changed the behaviours of treatment and comparison groups in different ways.</td>
<td>Score &quot;no information&quot; if the authors do not discuss potential bias due to monitoring of intervention.</td>
</tr>
<tr>
<td>4.3. Was an appropriate analysis used to estimate the effect of assignment to intervention?</td>
<td>Both intention-to-treat (ITT) analyses and modified intention to treat (mITT) analyses excluding participants with missing outcome data should be considered appropriate. (Missing outcome data are addressed in a separate domain). Both ‘as treated’ analyses (in which trial participants are grouped according to the intervention that they received, rather than according to their assigned intervention) and</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
naïve ‘per-protocol’ analyses (excluding trial participants who did not receive their assigned intervention) should be considered inappropriate. Analyses excluding eligible trial participants post-randomization should also be considered inappropriate, but post-randomization exclusions of ineligible participants (when eligibility was not confirmed until after randomization, and could not have been influenced by intervention group assignment) can be considered appropriate.

Score "Yes" if ITT or modified ITT analysis has been used, AND common analysis methods are used AND any covariates imbalanced at baseline are included as covariates in the analysis. Score "No" otherwise. You may also score no if they use an uncommon analysis methods, such as 3SLS, or if they analyse dichotomous outcome using continuous methods (e.g., using linear regression instead of using the more appropriate logit or probit models) or if the is baseline imbalance but those covariates are not included in the model.

If the analytical specification is not provided score "no information"

<table>
<thead>
<tr>
<th>Question 4.3 answer justification</th>
<th>Open answer</th>
<th>Justification for coding decision (Include a brief summary of justification for rating, mentioning your response to all sub questions, cite relevant pages).</th>
</tr>
</thead>
<tbody>
<tr>
<td>Deviations from intended interventions: Spill-overs, cross-overs, contamination and performance bias: was the study adequately protected against spill-overs, cross-overs, contamination, and performance bias?</td>
<td>(0) Low risk of bias (1) Medium risk of bias/Some concerns (2) High risk of bias</td>
<td>Utilize your answers for questions 4.1 to 4.3 to determine the overall score.</td>
</tr>
<tr>
<td>Deviations justification</td>
<td>Open answer</td>
<td>Justification for coding decision (Include a brief summary of justification for rating, mentioning your response to all sub questions, cite relevant pages). For example, intervention groups are geographically separated, authors use intention to treat estimation or instrumental variables to</td>
</tr>
</tbody>
</table>

4: Bias due to deviations from intended interventions - Justification
### 5. Outcome measurement bias

<table>
<thead>
<tr>
<th>Question 5.1 answer justification</th>
<th>Open answer</th>
<th>Justification for coding decision</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>5. Outcome measurement bias</strong></td>
<td><strong>5.1 Could the measurement of the outcome be different between the study arms?</strong></td>
<td><strong>5.1 Could the measurement of the outcome be different between the study arms?</strong></td>
</tr>
<tr>
<td></td>
<td>(1) Yes; (2) Probably yes; (3) Probably no; (4) No; (5) No information;</td>
<td>Comparable methods of outcome measurement (data collection) involve the same measurement methods and thresholds, used at comparable time points. Differences between intervention groups may arise because of ‘diagnostic detection bias’ in the context of passive collection of outcome data, or if an intervention involves additional visits to a healthcare provider, leading to additional opportunities for outcome events to be identified. Score ‘yes’ or ‘probably yes’ if the outcomes have been measured at different times periods for treatment and control arms or have used different methods of measurement. Score ‘no’ or ‘probably no’ otherwise. If no information is provided on the the time period and methods of data collection in treatment and control arms, then score ‘no information’</td>
</tr>
<tr>
<td><strong>5.2. Were the outcome assessors (enumerators) blinded to the intervention assignment?</strong></td>
<td>(1) Yes; (2) Probably yes; (3) Probably no; (4) No; (5) No information;</td>
<td><strong>5.2. Were the outcome assessors (enumerators) blinded to the intervention assignment?</strong> Score ‘yes’ or ‘probably yes’ if outcome assessors were blinded to intervention status and ‘no’ or ‘probably no’ otherwise. If no information is provided on blinding of the outcome assessors, then score ‘no information’. Note that for participant-reported outcomes, the outcome assessor is the study participant.</td>
</tr>
<tr>
<td><strong>5. Outcome measurement bias - Justification</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Question</td>
<td>Answer</td>
<td>Justification</td>
</tr>
<tr>
<td>----------</td>
<td>--------</td>
<td>---------------</td>
</tr>
<tr>
<td>5.3. If not blinded, could assessment of the outcome have been influenced by knowledge or administration of intervention received?</td>
<td>(1) Yes; (2) Probably yes; (3) Probably no; (4) No; (5) No information;</td>
<td>Knowledge or administration of the assigned intervention could influence participant-reported outcomes (such as level of pain, immunization outcomes using recall), observer-reported outcomes involving some judgement, and intervention provider decision outcomes. They are unlikely to influence observer-reported outcomes that do not involve judgement, for example all-cause mortality. For participant-reported outcomes, the assessment of outcome is potentially influenced by knowledge or administration of intervention received, leading to a judgement of at least ‘Some concerns’. Review authors will need to judge whether it is likely that participants’ reporting of the outcome was influenced by knowledge or administration of intervention received, in which case risk of bias is considered to be high. Score ‘yes’ or ‘probably yes’ if assessment of outcome could have been influenced by knowledge of intervention received (e.g., if ANY of the immunization coverage outcomes are assessed through caregiver recall) and ‘no’ or ‘probably no’ otherwise (e.g., if immunization coverage outcome is assessed by immunization card). If no information is available to make the assessment then score ‘no information’.</td>
</tr>
<tr>
<td>5. Outcome measurement bias - Justification</td>
<td>Outcome measurement bias: Was the study free from biases in outcome measurement?</td>
<td>(0) Low risk of bias (1) Medium risk of bias/Some concerns (2) High risk of bias</td>
</tr>
<tr>
<td>5. Outcome measurement bias - Justification</td>
<td>Outcome measurement justification</td>
<td>Open answer</td>
</tr>
<tr>
<td>6. Reporting bias</td>
<td>6.1. Is a pre-analysis plan or protocol available which provides sufficient detail?</td>
<td></td>
</tr>
<tr>
<td>------------------</td>
<td>-----------------------------------------------------------------------------</td>
<td></td>
</tr>
<tr>
<td>(1) Yes;</td>
<td>6.1. Is a pre-analysis plan or protocol available which provides sufficient detail?</td>
<td></td>
</tr>
<tr>
<td>(2) Probably yes;</td>
<td>Score 'yes' if they reference a preanalysis plan, and 'no' otherwise.</td>
<td></td>
</tr>
<tr>
<td>(3) Probably no;</td>
<td></td>
<td></td>
</tr>
<tr>
<td>(4) No;</td>
<td></td>
<td></td>
</tr>
<tr>
<td>(5) No information;</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>6. Reporting bias</th>
<th>6.1. Question 6.1 answer justification</th>
</tr>
</thead>
<tbody>
<tr>
<td>Open answer</td>
<td>Justification for coding decision</td>
</tr>
<tr>
<td></td>
<td>(Include a brief summary of justification for rating, mentioning your response to all sub questions, cite relevant pages).</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>6. Reporting bias</th>
<th>6.2. Question 6.2 answer justification</th>
</tr>
</thead>
<tbody>
<tr>
<td>Open answer</td>
<td>Justification for coding decision</td>
</tr>
<tr>
<td></td>
<td>(Include a brief summary of justification for rating, mentioning your response to all sub questions, cite relevant pages).</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>6. Reporting bias</th>
<th>6.3. Do reported results for the outcomes correspond to all intended analyses?</th>
</tr>
</thead>
<tbody>
<tr>
<td>(1) Yes;</td>
<td>6.3. Do reported results for the outcomes correspond to all intended analyses?</td>
</tr>
<tr>
<td>(2) Probably yes;</td>
<td>A particular outcome domain may be analysed in multiple ways. Examples include: unadjusted and adjusted models; final value vs change from baseline vs analysis of covariance; transformations of variables; different definitions of composite outcomes (e.g. ‘major adverse event’); conversion of continuously scaled outcome to categorical data with different cut-points; different sets of covariates for adjustment; and different strategies for dealing with missing data. Application of multiple methods generates multiple effect estimates for a specific outcome domain. If multiple estimates are generated but only one or a subset is reported on the basis of the results (e.g. statistical significance).</td>
</tr>
<tr>
<td>(3) Probably no;</td>
<td></td>
</tr>
<tr>
<td>(4) No;</td>
<td></td>
</tr>
<tr>
<td>(5) No information;</td>
<td></td>
</tr>
</tbody>
</table>
6. Reporting bias

<table>
<thead>
<tr>
<th>Question 6.3 answer justification</th>
<th>Open answer</th>
</tr>
</thead>
<tbody>
<tr>
<td>Justification for coding decision</td>
<td></td>
</tr>
</tbody>
</table>

(Include a brief summary of justification for rating, mentioning your response to all sub questions, cite relevant pages).

statistical significance), there is a high risk of bias in the fully reported result.

Answer ‘No’ or ‘Probably No’ if:
There is clear evidence (usually through examination of a trial protocol or statistical analysis plan) that a domain was analysed in multiple ways, but data for only one or a subset of analyses is fully reported (without justification), and the fully reported result is likely to have been selected on the basis of the results. Selection on the basis of the results arises from a desire for findings to be newsworthy, sufficiently noteworthy to merit publication, or to confirm a prior hypothesis. For example, trialists who have a preconception or vested interest in showing that an experimental intervention is beneficial may be inclined to selectively report analyses that are favourable to the experimental intervention. Score probably no if authors do not report both adjusted and unadjusted models.

Answer ‘Yes’ or ‘Probably Yes’ if:
There is clear evidence (usually through examination of a trial protocol or statistical analysis plan) that all reported results for the outcome domain correspond to all intended analyses.

or

Not all intended analyses have been reported but authors have convincingly justified the reasons for not doing so.

or

There is only one possible way in which the outcome domain can be analysed (hence there is no opportunity to select from multiple analyses).

or

Analyses are inconsistent across different reports on the same trial, but the trialists have provided the reason for the inconsistency and it is not related to the nature of the results.

Answer ‘No information’ if:
Analysis intentions are not available, or the analysis intentions are not reported in sufficient detail to enable an assessment, and there is more than one way in which the outcome domain could have been analysed.
<p>| 6. Reporting bias - Assessment | Analysis reporting: Was the study free from selective analysis reporting? | (0) Low risk of bias (1) Medium risk of bias (2) High risk of bias | Utilize your answers for questions 6.1 to 6.3 to determine the overall score. |
| 6. Reporting bias - Justification | Analysis reporting justification | Open answer | Justification for coding decision (Include a brief summary of justification for rating, mentioning your response to all sub questions, cite relevant pages). |
| 7. Other bias - Assessment | Other risks of bias Is the study free from other sources of bias? | 1 = Yes, 4 = No | - |
| 7. Other bias - Justification | Other bias justification | Open answer | Justification for coding decision (Include a brief summary of justification for rating, mentioning your response to all sub questions, cite relevant pages). For example, information is collected using a different survey instrument in different intervention groups; measurement of the intervention received is unclear. |
| 8. Blinding - observers - Assessment | Blinding of participants? | 1 = Yes 2 = No 8 = unclear 9 = N/A | If there is no information, code NO. If there is information but it is ambiguous, code UNCLEAR. |
| 8. Blinding - analysts - Assessment | Blinding of data analysts? | 1 = Yes 2 = No 8 = unclear 9 = N/A | If there is no information, code NO. If there is information but it is ambiguous, code UNCLEAR. |
| 8. Blinding - method(s) | Method(s) used to blind | Open answer (including describe method of placebo control) 9 = N/A | Describe method(s) used to blind |
| 9. External validity - Random Sampling | Was random sampling used? | 1 = Yes 2 = No | Was a random sampling method used (meaning everyone in the population had an equal chance of being selected for the study - note that this is different than random allocation to treatment versus control group)? Score yes if: (a) there is a sampling frame. The sampling frame is the actual list of individuals that the sample will be drawn from. Ideally, it should include the entire target population (and nobody who is not part of that population). AND (b) probability sampling is used (e.g., simple random, stratified random, cluster random, etc.) Score no if: individuals are selected based on non-random criteria, and not every individual has a chance of being included (e.g., a convenience sample, a purposive sample, a snowball sample, a voluntary response sample, etc.). |</p>
<table>
<thead>
<tr>
<th>9. External validity - Random Sampling Justification</th>
<th>Justification for answer to random sampling question</th>
<th>Open answer</th>
<th>Justification for coding decision (Include a brief summary of justification for rating, mentioning your response to all sub questions, cite relevant pages). Describe the sampling technique used.</th>
</tr>
</thead>
<tbody>
<tr>
<td>9. External validity - Assessment</td>
<td>External validity</td>
<td>Open answer</td>
<td>a) What do authors say about external validity? Note any additional information related to generalizability.</td>
</tr>
</tbody>
</table>
Qualitative risk of bias tool

Critical appraisal of qualitative studies tool

This tool provides specific questions you should answer with regards to the study being appraised. Every question has three possible responses, which correspond to ‘strong’, ‘weak’, or ‘none’ – these are phrased slightly differently depending on the question. This process is subjective and relies on a basic level of familiarity with qualitative research methods. However, to further structure the assessment, each question is accompanied by a description of criteria you should consider before making your choice.

While attempting to answer a question, you may find that a study contains many of the relevant elements, but not all. Give this paper the benefit of the doubt – if the element being assessed is strong except for one minor element, choose the ‘strong’ option.

There is room for comments after each question. You are encouraged to utilize this space, though only the multiple answer response is mandatory. Use the space to document your thoughts if you feel uneasy about the choice made, or if you have any ideas which would help as we refine and revise the assessment tool. There is also room at the end for any additional notes on the paper as a whole.

The questions are arranged according to the order in which a standard academic article is arranged: Introduction, Methodology, Results, Discussion. This will allow you to answer the questions as you go over each section. That said, some documents will not follow this structure, and some academic articles are more fluid about where certain descriptions are located within the text. After you read through the article, go over all the questions once more to see whether any elements you thought were missing actually showed up in an unexpected place.

* Required

Assessor: *

Paper Title: *

Authors: *

Year: *

Introduction

1. Is the research aim clearly stated? *

Look for this in either the abstract or the introduction. In the best studies, this will be an explicit formulation of the research aims/questions. There may be several. Other studies do not formally state the research aims as such, but they are clearly evident from the text. Note: you might find it useful to write down these research aims, as many of the other questions relate to them.

• Yes, a strong statement
• Yes, but an unclear or weak statement
• No

Question 1 – Notes:

2. Is there a description of the context in which the study takes place? *


The study should make some reference to the geographical, temporal, or societal greater context within which the study topic is located. A study within a well-established area of research could likely be situated quite specifically. An exploratory study might cover a topic where there is little descriptive context to draw upon.

- Yes, a strong description
- Yes, but a weak description
- No

Question 2 – Notes:

3. Is there a clear link to relevant literature? *

Literature should be cited not just in general, but in specific relation to the topic of the paper. In assessing whether the literature is relevant and sufficient, consider: Are most of the cited papers about the same area or topic as the study? Are both qualitative and quantitative studies cited? Are there at least some papers published within the last 10 years?

- Yes, a clear strong link
- Yes, but it could be improved
- No

Question 3 – Notes:

4. Is there a clear link to theory? *

This can be achieved in many ways, and while it will often be in the introduction, it may show up elsewhere, so keep your eyes peeled. This requires a reference to a specific theory. There are two main ways this could happen: A reference could be made to an established theoretical framework (from the social sciences, from public health, etc.) which has been used to investigate similar questions, or which the authors think might apply in this case. The second possibility is that the authors describe a theoretical framework they have created or adapted themselves which they used in structuring their study, or that they refer to in considering their results.

- Yes, a strong one
- Yes, but a weak one
- No

Question 4 – Notes:

Methodology - Sampling

5. Is there a description of the sampling procedure? *

Are there details about how sampling was actually conducted? A helpful way to consider this is whether you would be able to conduct this sampling yourself if you were to replicate the study. Every sampling approach has its own logic, but even for the more straightforward qualitative sampling, there should be some mention of this. Note: This
question does not evaluate whether this sampling is actually appropriate (see question 6). It only asks whether we have a clear idea of the process of sampling, at all.

- Yes, a strong description
- Yes, but a partial description
- No

Question 5 – Notes:

6. Is the sampling strategy appropriate for the aims of the research? *

This question asks whether the sampling strategy was the right for choosing the participants/locations/etc. that could yield relevant information. If there is a strong claim of generalizability, was the sampling appropriate for achieving this? If the study attempts to present the views of a group, did sampling actually capture the people the study needed to hear from? Are there any groups which appear to be missing but clearly should have been spoken to in order to meet the research aims? Has saturation been achieved, and can we tell how this was established? Note that this is not a question of whether the method of collection (interview, focus group, etc.) would have yielded strong data – just whether the people recruited to the study were the right people.

- Yes, very appropriate
- Yes, but not completely appropriate
- No

Question 6 – Notes:

7. Are sample characteristics sufficiently reported? *

The basic characteristics are often gender, location, ethnicity, and other demographic variables, but consider whether there are other sample characteristics specific to the question which should be reported. This might be clear by this point, but you might also want to look ahead and see whether the paper discusses any relevant characteristics that you would want to know about the sample in general (or you could return to this at the end).

- Yes, there's a sufficient description
- There's some description, but not enough
- No

Question 7 – Notes:

Methodology - Data Collection

8. Is it clear how data were collected? *

Each method requires its own information. If interviews were used, is there detail on locations, presence of others, length of interview, questions asked, and other such details? For participant observation, is there a description of the
time spent in the field and the activities partaken in? For focus groups, is there detail on the number of participants, questions asked and setting?

- Yes, data collection is clearly described
- Yes, but more details are needed
- No

Question 8 – Notes:

9. Are the methods of data recording reported? *

This could be as simple as ‘the audio of interviews was taped’, ‘focus group summaries were written during the discussion’, or ‘interviews were transcribed, translated, and the translation was then verified’. Any of these would merit at least a ‘weak’ response, and more detail would be considered ‘strong’.

- Yes, and reported well
- Yes, but only briefly mentioned
- No

Question 9 – Notes:

10. Did the collection of the data address the research aims? *

Given the data needed to answer the question, is the method chosen actually well suited to obtaining such data? If we have knowledge of the questions in an interview guide or a focus group guide, are all the relevant subject areas covered? This isn’t about whether there might be some problems in the collection (a bad interviewer could screw up even the best interview guide) but whether the method of data collection, if employed properly, would actually capture relevant data.

- Yes, completely
- Yes, but not entirely
- No

Question 10 – Notes:

Methodology - Analysis

11. Are the methods of analysis explicitly stated? *

This should include a basic statement as to the approach used (thematic analysis, grounded theory, etc.) but also a breakdown of the steps taken (different rounds of coding, intermittent checks for quality, revisions, etc.)

- Yes, and stated well
- Yes, but more detail is needed
- No
Question 11 – Notes:

12. Were there any inbuilt checks to assure the quality of the analysis? *

Some but not all of the following could be present. Was there more than one researcher involved in the analysis? Was there independent peer review during the analysis? Were inter-coder or intra-coder reliability checks conducted? Did the team meet regularly to discuss the analytic process? Did the researcher keep a journal of their thoughts and decisions?

• Yes, there were clear, strong checks throughout
• Yes, but they were weak checks
• No

Question 12 – Notes:

13. Was there reflection on bias and positionality? *

Have the authors considered their potential biases in relation to the study? This should be a specific statement, and there should be mention of steps that were taken to address this, even if as basic as taking the time to write reflective notes or engaging in a group discussion of the issues.

• Yes, a clear reflection with detail provided
• Yes, but there's only a brief mention of this
• No

Question 13 – Notes:

Methodology - Researchers

14. Are there any details about the people who conducted the sampling, data collection, and analysis?

Not much is required, but look out for mentions of the training these people had, their positions/titles, their experience, etc.

• Yes, plenty
• Yes, a little
• No

Question 14 – Notes:

Results

15. Is there enough data to support the claims? *
Variable, but consider whether major claims/themes are supported by more than a single quote or other type of original data. Look at whether the data presented is contextualised, and whether it demonstrates the claim that the authors suggest it does.

- Yes, there's plenty of clear, contextualised, relevant data
- There's some data, but it's weak
- No - there's little or no data presented that is supports the claims.

Question 15 – Notes:

16. Are diverse viewpoints considered? *

At a basic level, we need to know who speakers are if they are quoted. This could be as minimal as ('22-year-old', or 'Mother, Group A'). Same goes for summaries with no direct quotes – it should be clear whose viewpoint the data comes from. The majority of the data presented will demonstrate the main themes/arguments, but no study is without some dissenting voices or unusual opinions. Are these ever mentioned? If there were many participants, but it appears that most quotes come from one or two individuals, there should be a reasonable explanation for why this is the case – and what it might mean for the study.

- Yes, thoroughly
- Yes, only occasionally
- No

Question 16 – Notes:

17. Is there evidence which addresses every research aim? *

Look back to the research aims. Is there data relating to each one of them? If not, is there a clear reason for this? If a study ended up not finding data relevant to every aim, but acknowledges that this is the case and tries to account for the reason, this should still be marked as a yes. A 'null' result is still a result - sometimes collection and analysis can be excellent, yet no clear answer is found. If this appears to be the case, the study can still be considered 'strong'.

- Yes, every aim has some related evidence
- Yes, but some aims have distinctly less attending to them
- No

Question 17 – Notes:

Discussion

18. Has the question been answered? *

A simple test is to go back to the research aims, formulate them as questions (if they’re not already established as such) and see whether the authors have provided you with answers that are reasonable given the data presented. It should be clear how the data presented is appropriate and convincing evidence for the conclusion. In a strong paper,
the authors will make a clear and logical statement of how the results provide an answer to the question. A weaker paper might not cover all the questions, or might suggest questionable links between the data in the results and the conclusions reached.

- Yes, with a clear, logical, thorough answer
- Yes, but not well
- No

Question 18 – Notes:

19. Are relevant literature, theory, or practice discussed in relation to the results? *

A good paper doesn’t need to discuss all of these, but it should relate to at least one. Did it match up or challenge a theory? Does it have implications for practice? Are the results now situated within the broader literature?

- Yes, and discussed well
- Yes, but only briefly or generally discussed
- No

Question 19 – Notes:

20. Has there been any triangulation? *

This takes many forms, but the basic question is – do the authors present any other sources of data which support or contradict their claims? These could be quantitative/qualitative/mixed studies, some form of evaluation of the conclusions by the participants themselves, comparison to other data recorded, etc. This could be in the results section or the discussion, so have another look.

- Yes, strong triangulation
- Yes, weak triangulation
- No

Question 20 – Notes:

21. Are weaknesses considered? *

Mentioning them is a good step, but a strong paper will also argue how these weaknesses were attended to.

- Yes, major weaknesses considered and attended to
- Yes, but weaknesses are mentioned without much discussion
- No

Question 21 – Notes:
Ethics

22. Have ethical issues been taken into consideration? *

At the very basic level, there should be a description of informed consent. Also look for approval from an IRB (or other ethics committee as applicable), and whether issues of benefit/harm and confidentiality/secrecy were discussed with participants.

• Yes
• No

Question 22 – Notes:

General comments and thoughts:

Remember to look back at the questions before submitting.

Having read the paper as a whole, and having considered each question separately, you might now also have a clearer idea as to the answers to some questions. Have another look, particularly for questions where you noted some thoughts or concerns. You may want return to question 7 in particular, as you may now have a clearer idea of what sample characteristics should have been reported.
Cost evidence risk of bias tool

<table>
<thead>
<tr>
<th>S.No.</th>
<th>Questions</th>
<th>Response</th>
</tr>
</thead>
<tbody>
<tr>
<td>1a.</td>
<td>Is the form of economic evaluation clearly stated? Indicate (CBA/CEA/CUA/cost-minimization, or innovative approaches: Social Return on Investment (SROI), Multi-Criteria Appraisal (MCA). Response options [CBA, CEA, CUA, SROI, MCA, or N/A if only &quot;total cost&quot;, also ok to insert descriptive results that do not conform, e.g. “cost transfer ratio”]</td>
<td>Yes, No</td>
</tr>
<tr>
<td>1b.</td>
<td>Is the perspective of the costing stated?</td>
<td>Yes, No</td>
</tr>
<tr>
<td>1d.</td>
<td>Cost data sources?</td>
<td>Yes, No</td>
</tr>
<tr>
<td>1e.1</td>
<td>Are unit costs reported (in Table)</td>
<td>Yes, No</td>
</tr>
<tr>
<td>1e.2</td>
<td>Are cost ingredients listed?</td>
<td>Yes, No</td>
</tr>
<tr>
<td>1e.3</td>
<td>Total cost description</td>
<td>Yes, No</td>
</tr>
<tr>
<td>1f.</td>
<td>Total cost reported?</td>
<td>not reported, Value if reported</td>
</tr>
<tr>
<td>1h.</td>
<td>Was cost per child immunised reported?</td>
<td>Yes, No, If yes, Cost</td>
</tr>
<tr>
<td>1j.</td>
<td>What exchange rate was used for currency adjustments? [e.g. 1.2 USD to 1 Euro]</td>
<td>value if given, Not reported</td>
</tr>
<tr>
<td>1j.</td>
<td>Exch rate year (estimated)</td>
<td>Year</td>
</tr>
<tr>
<td>1k.</td>
<td>Is time horizon for costs clearly stated? [Operationalise - was a base year of the costing reported?</td>
<td>Yes, No, If yes: year</td>
</tr>
<tr>
<td>7a.</td>
<td>Sensitivity analysis? [yes/ no]</td>
<td>n/a, Description if yes</td>
</tr>
</tbody>
</table>
Appendix 8: Risk of bias in the included studies

Quantitative risk of bias

Risk of bias in Randomised Controlled Trials

Of the included studies with experimental designs (k = 31) the majority (k = 23) have been identified as being at high risk of bias, and only two were rated as having a low risk of bias. Six studies were assessed as having some concerns. Of the six potential causes of bias that were analysed, outcome measurement bias and deviations from intended interventions were the most commonly documented issues. The most common pitfall related to outcome measurement bias was the use of caregiver reported (self-reported) measures of vaccinations received by a child in the absence and/or incompleteness of immunisation cards, which is common in L&MICs and is often influenced by the intervention itself. One such study stated that “The calculation of FIC coverage was based on data from the child’s immunization card and/or the mother’s recall. Mother’s recall was taken in situations in which the card or specific data points on the card were missing.”, (Gurley 2020). Relatively, Demilew, (2021) reported that their outcome measures “…could be subject to recall bias or social desirability bias differentially by treatment group…”. Deviations from intended interventions had to do with issues such as potential spillover related to geography (e.g. in Siddiqi 2020, centers were contiguously located) or transferable vouchers (e.g. Morris 2004). Other sources of bias in included RCT studies are reporting bias, and bias in the randomisation process. Supplementary Figure 2a illustrates the frequencies with which each cause of bias was identified across the 31 included RCTs, while Supplementary Figure 2b shows the risk of bias assessments for each individual study and each risk of bias category for RCT studies.

Supplementary Figure 2a: Frequencies of Bias in RCTs

<table>
<thead>
<tr>
<th>Sources of Bias</th>
<th>Low risk of bias</th>
<th>Some concerns</th>
<th>High risk of bias</th>
</tr>
</thead>
<tbody>
<tr>
<td>Randomisation process</td>
<td>19</td>
<td>8</td>
<td>4</td>
</tr>
<tr>
<td>Deviations from intended interventions</td>
<td>19</td>
<td>3</td>
<td>9</td>
</tr>
<tr>
<td>Missing outcome data</td>
<td>23</td>
<td>4</td>
<td>4</td>
</tr>
<tr>
<td>Outcome measurement bias</td>
<td>11</td>
<td>2</td>
<td>18</td>
</tr>
<tr>
<td>Reporting bias</td>
<td>13</td>
<td>15</td>
<td>3</td>
</tr>
</tbody>
</table>

BMJ Publishing Group Limited (BMJ) disclaims all liability and responsibility arising from any reliance placed on this supplemental material which has been supplied by the author(s).
Supplementary Figure 2b: Risk of bias assessments for individual RCT studies

<table>
<thead>
<tr>
<th>Study first author</th>
<th>Assignment mechanism: Was the allocation or identification mechanism random or as good as random?</th>
<th>Unit of analysis: Is unit of analysis in cluster allocation addressed in standard error calculation?</th>
<th>Bias due to missing outcome data</th>
<th>Was the study adequately protected against spill-overs, cross-overs, contamination, and performance bias?</th>
<th>Outcome measurement bias: Was the study free from biases in outcome measurement?</th>
<th>Analysis reporting: Was the study free from selective analysis reporting?</th>
<th>Overall Risk of Bias Assessment</th>
</tr>
</thead>
<tbody>
<tr>
<td>Alhassan, 2019</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Andersson, 2009</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Arifeen, 2009</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Banerjee, 2010</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Banerjee, 2020</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Björkman, 2009</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Bolam, 1998</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Borkum, 2014</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Demilew, 2020</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Domek, 2019</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Engineer, 2016</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gibson, 2017</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gurley, 2020</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Johri, 2018</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lee, 2015</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Modi, 2019</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mohanan, 2020</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>More, 2012</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>More, 2017</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Morris, 2004</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Murthy, 2019</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Nagar, 2018</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Nagar, 2020</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Olken, 2014</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Oyo-Ita, 2020</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pramanik, 2020</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Rahman, 2008</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Robertson, 2013</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Seth, 2018</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Siddiqi, 2020</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Webster, 2019</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Note: Green cells signify low risk of bias, yellow cells signify some concerns related to risk of bias, and red cells signify high risk of bias.
Bias in Quasi-Experimental Designs

Of the 30 included quasi-experimental studies, the majority (k = 27) present a high risk for bias, two were assessed as low risk of bias and one study was assessed as having some concerns. Like the experimental studies the most common identified causes of bias amongst quasi-experimental designs were issues related to outcome measurement bias and deviations from intended interventions. This is no surprise as self-reporting was again a common concern (e.g. Admassie, 2009 and Findley, 2013). Supplementary Figure 3a illustrates the frequencies of other biases in our included QEDs, while Supplementary Figure 3b shows the risk of bias assessments for each individual study and each risk of bias category for QED studies.

### Supplementary Figure 3a: Frequencies of Bias in QEDs

<table>
<thead>
<tr>
<th>Sources of Bias</th>
<th>Low risk of bias</th>
<th>Some concerns</th>
<th>High risk of bias</th>
</tr>
</thead>
<tbody>
<tr>
<td>Confounding</td>
<td>9</td>
<td>7</td>
<td>14</td>
</tr>
<tr>
<td>Deviations from intended interventions</td>
<td>7</td>
<td>7</td>
<td>15</td>
</tr>
<tr>
<td>Missing outcome data</td>
<td>23</td>
<td></td>
<td>3, 4</td>
</tr>
<tr>
<td>Outcome measurement bias</td>
<td>9</td>
<td>2</td>
<td>19</td>
</tr>
<tr>
<td>Reporting bias</td>
<td>11</td>
<td>8</td>
<td>11</td>
</tr>
</tbody>
</table>

### Supplementary Figure 3b: Risk of bias assessments for individual QED studies

<table>
<thead>
<tr>
<th>Study first author</th>
<th>Was the identification method free from any sources of bias</th>
<th>Bias due to missing outcome data</th>
<th>Was the study adequately protected against spillovers, crossovers, contamination, and performance bias?</th>
<th>Outcome measurement bias: Was the study free from biases in outcome measurement?</th>
<th>Analysis reporting: Was the study free from selective analysis reporting?</th>
<th>Overall Risk of Bias Assessment</th>
</tr>
</thead>
<tbody>
<tr>
<td>Adamu, 2019</td>
<td>Yellow</td>
<td>Green</td>
<td>Red</td>
<td>Yellow</td>
<td>Red</td>
<td>Red</td>
</tr>
<tr>
<td>Admassie, 2009</td>
<td>Yellow</td>
<td>Green</td>
<td>Red</td>
<td>Red</td>
<td>Yellow</td>
<td>Red</td>
</tr>
<tr>
<td>Assegaai, 2018</td>
<td>Yellow</td>
<td>Green</td>
<td>Red</td>
<td>Red</td>
<td>Yellow</td>
<td>Red</td>
</tr>
<tr>
<td>Banwat, 2015</td>
<td>Yellow</td>
<td>Green</td>
<td>Red</td>
<td>Red</td>
<td>Yellow</td>
<td>Red</td>
</tr>
<tr>
<td>Biemba, 2016</td>
<td>Yellow</td>
<td>Green</td>
<td>Red</td>
<td>Red</td>
<td>Yellow</td>
<td>Red</td>
</tr>
<tr>
<td>Calderon-Ortiz, 1996</td>
<td>Yellow</td>
<td>Green</td>
<td>Red</td>
<td>Red</td>
<td>Yellow</td>
<td>Red</td>
</tr>
</tbody>
</table>
Qualitative risk of bias
Risk of bias assessments were conducted on 47 qualitative papers. Papers were scored as absent, weak, or present on 12 key elements and received corresponding scores of zero, one, or two for these ratings. Most (27) papers were missing at least some key elements, resulting in them receiving a zero value for these elements in their quality assessment scores (Supplementary Table 1; Supplementary Figure 4). However, 17 received quality assessment scores over 20, indicating that they received a value of two, or a “strong” rating, for most key elements (Supplementary Figure 5). The most common key elements to be missing were descriptions of sample characteristics and the analytic methods, with 17 studies failing to report on each of these (Supplementary Figure 5). The research aim was the most common key element to be stated strongly, with 39 papers clearly stating their aims.

Supplementary Table 1: Distribution of papers by strength of key elements

<table>
<thead>
<tr>
<th>Number</th>
</tr>
</thead>
<tbody>
<tr>
<td>One or more key elements absent</td>
</tr>
<tr>
<td>One or more key elements weak</td>
</tr>
</tbody>
</table>
Supplementary Figure 5: Distribution of qualitative assessment scores*

Risk of bias in cost effectiveness estimates

The risk of bias analysis of the cost and cost-effectiveness analysis assessed the quality of underlying cost data, reporting, and analysis using information that was reported in the immunization studies. We assessed risk of bias along six primary dimensions which were adapted from a combination of tools, including: Doocy and Tappis (2017); Campbell Collaboration Economic Methods Policy Brief (Shemilt et al.,2008); and Methods for the Economic Evaluation of Health Care Programmes, (Drummond et al. 2015), that were adapted for cost analyses carried out in conjunction with impact evaluation studies of global development interventions.

*Papers received scores of 0, 1 or 2 reflecting the strength of reporting on 12 key elements in the risk of bias assessment tool (Appendix E).
This risk of bias tool specifically assesses the bias that arises from the collection and reporting of cost analysis in conjunction with impact evaluation studies of global development interventions. These studies often do not incorporate cost analyses. Indeed recent estimates suggest just 15-18% of impact analyses include any kind of cost analysis (Brown and Tanner 2019). This under-reporting of cost in conjunction with impact evaluations leads to very small samples from which to draw inferences about the cost, and cost-effectiveness of development interventions. Moreover, often when estimates are included with impact evaluations, the quality of the underlying data is low. Cost may have been added as an “after-thought”, rather than planned for in advance or treated as a research endeavor. Data sources for cost information may not be well-documented or may have been estimated using ‘back-of-the-envelope’ techniques (this phrase literally is used in write-ups of cost methods). There is often insufficient detail in reporting of cost to assess the quality of estimates and data and basic robustness checks of the analysis are very infrequently performed.

Finally, this tool specifically considers the elements of cost and effectiveness that were extracted for the analysis of incremental cost effectiveness. Our particular measure estimates the non-vaccine cost per dose of interventions to increase absolute immunization coverage by one percent, using available data extracted from the evaluations wherever possible, or by following the steps and calculations as outlined in Ozawa et al. (2018). The specific inputs to this analysis include estimates of: total intervention cost (net of vaccine cost); the number of vaccine doses provided to each child; the endline proportion of children that received immunisations in the treatment and control groups; and the cost per child immunised.

Dimensions of potential bias in costs reported and cost-effectiveness estimates

**Planned, organized, cost analysis.** A common challenge of cost analysis is a lack of planning which often can lead to poor underlying data quality. For example, cost estimates are subject to recall bias when assessed long after the program and evaluation are completed. Therefore, we examine three indicators of an organized or planned cost analysis. Specifically, we look for a clear description of the form of economic evaluation, and a description of the method used. We also look for a clear statement of analytical perspective--which is the choice of which actor has standing in the costing -- this may be the donor, the implementing partner, or perspective of the costing may be from a societal point of view. The analytical perspective is important because it determines whose costs and benefits will be counted in the costing. It may, for example, directly impact how the constituent components of total cost are counted. Although we considered a standard ROB question which asks: ‘Is a well-defined research question posed in answerable form?’, we find that impact evaluation studies do not pose research questions in the expected form. The three questions, given equal weight in the risk of bias assessment, were:

a. Is the form of economic evaluation clearly stated?
b. Is the perspective of the costing stated?  
c. Was a method of costing described?

**Quality data sources.** The highest quality data for assessing the cost of interventions in low-and middle-income settings often are drawn from the expenditure reports or accounting statements of the program implementers. Expenditures are better quality because they represent actual, rather than planned expenses and they often are subject to audit, making them more reliable as compared with cost data taken from program budgets (Levin et al. 2018).

d. What is the quality of the primary data sources used for the cost estimates?

**Descriptive, detailed cost information.** The quality of descriptive detail on costs allows judgement into cost components and whether they align with intervention activities. Since the three criteria tend to be positively correlated, we give equal weight to the three indicators of detailed cost information. Specifically, we assessed the descriptive detail of reported costs:

e. Whether costs are reported by ingredients (or input or resource),  
f. Whether unit costs are reported, and
g. Whether the information is presented in an organized cost table

We consider the **quality of cost estimates** that were key inputs to the cost synthesis, specifically:

h. Total cost: What is the quality of the specific data components of total cost? Since total cost is a key input to the analysis we examine how it was reported. For example, some studies say simply, the total cost of the program was USD $3 Million. In the absence of descriptive information, i.e. the elements that comprise total cost, it is difficult to assess the quality or reliability of the total cost estimate.

i. Vaccine cost: Can we tell if vaccines were excluded? Since we need to exclude vaccine cost, it must be excluded or reported separately from total cost to be valid in this study.

j. Cost per child immunized: If cost per child immunized was not directly estimated and reported, we take total cost divided by the number of children in the treatment group multiplied by the proportion of children that were immunized. If the quality of underlying data are poor, this estimate may be biased – with unclear magnitude and direction.

k. Number of vaccines. Was the number of vaccines reported from observation of study participants, or was it reported per protocol? When we estimate cost per vaccine dose, the cost estimate will be biased downwards (less costly per dose) if the study does not report the actual number of doses administered. This is because some children may already have received a part of the vaccine protocol before enrolling in the study.

Are **key cost details** reported which allow us to adjust for time and currency differences? In this assessment of bias risk, the responses are given equal weight since we need both elements to accurately adjust for time and currency differences.

l. Are details of inflation and currency conversion clearly stated?

m. Is the time horizon for costs clearly stated?

We consider the **quality of cost analysis**, specifically:

n. Is sensitivity analysis conducted?

**Results**

We developed evaluation criteria and applied the protocol described above to the 22 immunisation evaluations that reported cost information. The results of this analysis are summarized below in Supplementary Figure 6, below.

Note: green cells indicate that clear information was presented in the report to assess the risk of bias; yellow cells indicate incomplete information; and red cell indicate we found very little or no information to assess bias risk.
Supplementary Figure 6: Risk of bias assessment of cost and cost-effectiveness estimates used in the analysis

- **Just over half of the included evaluations appear to have carried out a planned, organized, cost analysis.** Thirteen of 22 evaluations clearly state the form of economic evaluation (i.e. cost-effectiveness analysis); ten studies report the perspective of the costing, which is key for judging the correct inclusion and exclusion criteria for the components of a total cost estimate; and eleven of 22 evaluations describe the method of costing that was used to collect cost data (i.e. the ingredients method).

- **Indications of the quality of underlying cost data are mixed.** Nine of the 22 evaluations used expenditure reports to generate cost estimates, two used budgets and the remaining 11 evaluations provided no information on the provenience of the underlying cost data that was used in the analysis.

- **The quality of the descriptive detail on reported costs was mixed.** Just over half of all evaluations (14 of 22) provided thorough, descriptive information on costs; two evaluations provided some descriptive information, e.g. a breakdown of key unit costs; and six evaluations gave very minimal or no descriptive information on costs.

- **In 17 of 22 evaluations, we have high confidence that total cost excludes vaccines costs.** A majority of evaluations reported an estimate of total cost (20 of 22 evaluations). In 16 of the 20 total cost estimates, there were clear indications that vaccine costs had been excluded from the estimate, and in one case, vaccine costs were included but reported separately so that we could subtract vaccine cost from total cost to derive the comparable total cost. It was not possible to tell if vaccines were excluded from the total cost estimates of five evaluations.

- **The cost per immunized child** – was only reported by the authors in five of 22 evaluations and for the remaining 16 was estimated by the authors to compile Supplementary Table 22.

- **The number of vaccine doses received per treated child was reported in only four evaluations;** in 15 studies, vaccine doses per child were estimated based on information “per protocol” which does not account for children’s partial vaccination status at the point of enrollment in the study.

- **Fewer than half of the evaluations included the key information needed to adjust for time, currency, inflation or base year differences in the timing of expenditures.** This lack of reporting makes it very difficult to have complete confidence in the comparability of the cost estimates we generated.

- **Only three evaluations reported any kind of sensitivity analysis,** an indication of analytical robustness.
Appendix 9: Quantitative results – all community engagement interventions

Appendix A presents supplementary data for analyses presented in the main body of the manuscript for the primary outcomes of community engagement interventions, including funnel plots of publication bias and outlier analyses. It also includes a full presentation of the supplementary outcomes that were precluded from the main body of the manuscript due to word limitations.

Primary outcomes

Full immunisation

As a robustness check, we used robust variance analysis and included all dependent effects in the analysis, totaling 53 effects from the same 28 studies ($df = 25.5$). The overall average effect was slightly smaller but still significant ($\hat{\mu} = 0.11$ [95% CI: 0.04 to 0.18], $p = .002$). Sensitivity analyses show the effect to be sensitive to all values of Rho. The rank correlation test indicated funnel plot asymmetry ($p = 0.03$) but the regression test did not ($p = 0.57$; see Supplementary Figure 7). A trim and fill analysis indicated an identical effect size (see Supplementary Figure 7b).

Supplementary Figure 7. Funnel plot for studies examining the effect of community engagement interventions on full immunisation
Supplementary Figure 7b. Trim and fill plot for studies examining the effect of community engagement interventions on full immunisation

**DPT3**

According to the $Q$-test, the true outcomes appear to be heterogeneous ($Q(21) = 90.43, p < 0.01, \hat{t}^2 = 0.01, I^2 = 76.78\%$). The rank correlation test indicated funnel plot asymmetry ($p = 0.04$) but not the regression test ($p = 0.06$), but trim and fill analyses indicated an identical effect size (Supplementary Figures 8 & 9). An examination of the studentized residuals revealed that none of the studies had a value larger than ±3.05 and hence there was no indication of outliers in the context of this model. According to the Cook’s distances, none of the studies could be considered to be overly influential. Leave one out analyses indicated that no single study was driving this result.

Publication year was a significant moderator such that each additional year reduced the size of the effect by .014 standard deviation units ($\hat{B} = -0.014, p = 0.019$ [95% CI: −0.03 to −0.002]). In other words, new studies have found smaller effects. There were no other significant moderators in the context of this model.

Supplementary Figure 8. Funnel plot for studies examining the effect of community engagement interventions on DPT3 vaccination
As a robustness check, we used robust variance analysis and included all dependent effects in the analysis, totaling 36 effects from the same 22 studies ($df = 18.2$). The overall average effect was slightly smaller but still significant ($\hat{\mu} = 0.10$ [95% CI: 0.05 to 0.15], $p < .001$). Sensitivity analyses show the effect to be sensitive to all values of Rho.

**Measles**

An examination of the studentized residuals revealed that one study (Sankar 2013) had a value larger than $\pm 3.02$ and may be a potential outlier in the context of this model. Indeed, sensitivity analysis leaving out Sankar (2013) would result in an increase in the average effect ($\hat{\mu} = 0.08$ [95% CI: 0.05 to 0.12], and it was still statistically
significant ($z = 4.44, p < 0.001$). Leave one out analyses indicated that no other study was driving this result. According to the Cook’s distances, none of the studies could be considered to be overly influential. A funnel plot of the estimates is shown in Figure A6. Both the rank correlation and the regression test indicated potential funnel plot asymmetry ($p = 0.05$ and $p = 0.02$, respectively). The trim and fill analysis (Supplementary Figure 13) estimates a smaller and non-significant effect ($\hat{\mu} = 0.02 [95\% CI: -0.02 to 0.06], p = .36$ Supplementary Figure 11 presents the results when low quality studies are removed.

![Forest plot showing the observed outcomes and the estimate of the random-effects model for the impact of community engagement interventions on measles vaccination when low quality studies are removed](image1)

**Supplementary Figure 11.** Forest plot showing the observed outcomes and the estimate of the random-effects model for the impact of community engagement interventions on measles vaccination when low quality studies are removed

![Funnel plot for studies examining the effect of community engagement interventions on measles vaccination](image2)

**Supplementary Figure 12.** Funnel plot for studies examining the effect of community engagement interventions on measles vaccination
Supplementary Figure 13. Trim and fill funnel plot for studies examining the effect of community engagement interventions on measles vaccination

As a robustness check, we used robust variance analysis and included all dependent effects in the analysis, totaling 32 effects from the same 20 studies (\(df = 15.5\)). The overall average effect was slightly smaller but still significant (\(\hat{\mu} = 0.06\) [95% CI: 0.01 to 0.11], \(p = .03\)). Sensitivity analyses show the effect to be sensitive to all values of Rho.

Vaccination timeliness: Full immunisation schedule

We included a total of \(k = 5\) studies in the analysis. The estimated average outcome was \(\hat{\mu} = 0.15\) ([95% CI: 0.07 to 0.24], \(z = 3.41, p < 0.001\); Supplementary Figure 14), indicating a small but significant benefit to the treated group compared to the control group. According to the \(Q\)-test, there was no significant amount of heterogeneity in the true outcomes (\(Q(4) = 4.43, p = 0.35, \tau^2 = 0.00, I^2 = 9.66\%\)). An examination of the studentized residuals revealed that none of the studies had a value larger than ±2.58 and hence there was no indication of outliers in the context of this model. This was confirmed by a leave-one-out analysis. According to the Cook’s distances, none of the studies could be considered to be overly influential. With only one study not assessed as high risk of bias, we were unable to conduct sensitivity analyses by study quality. With no significant heterogeneity, we did not test for moderations. There were no dependent effects, so we could not use RVE as a robustness check in this case.
Supplementary Figure 14. Forest plot showing the observed outcomes and the estimate of the random-effects model for the impact of community engagement interventions on timeliness of full childhood immunisation

Vaccination timeliness: DPT3

A total of \( k = 7 \) studies examined the relationship between community engagement interventions and timeliness of DPT3 vaccination. The estimated average outcome was \( \hat{\mu} = 0.09 \) ([95% CI: 0.03 to 0.14]), \( z = 3.00, p < 0.01 \), indicating a small but significant benefit to the intervention group compared to the control group (Supplementary Figure 15). According to the \( Q \)-test, there was no significant amount of heterogeneity in the true outcomes (\( Q(6) = 5.09, p = 0.53, I^2 = 0.00\% \)). An examination of the studentized residuals revealed that none of the studies had a value larger than \( \pm 2.69 \) and hence there was no indication of outliers in the context of this model. This was confirmed by a leave-one-out analysis. According to the Cook’s distances, none of the studies could be considered to be overly influential. With no heterogeneity, moderator analyses were not appropriate. There were no dependent effects, so we could not use RVE as a robustness check in this case. When low quality studies are removed, only two studies remain, and the estimated average outcome increases to \( \hat{\mu} = 0.11 \) ([95% CI: −0.05 to 0.27]), but becomes non-significant \( z = 1.32, p = 0.19 \). However, with only two studies this should be interpreted with caution.
Supplementary Figure 15. Forest plot showing the observed outcomes and the estimate of the random-effects model for the impact of community engagement interventions on the timeliness of DPT3 vaccination.

Vaccination timeliness: Measles

Only $k = 2$ studies using community engagement interventions reported on the timeliness of measles vaccinations. The estimated average outcome was $\hat{\mu} = 0.23$ ([95% CI: 0.14 to 0.32], $z = 5.06$, $p < 0.001$; see Supplementary Figure 16), indicating a small but significant benefit to the treated group compared to the control group. According to the $Q$-test, there was no significant amount of heterogeneity in the true outcomes ($Q(1) = 0.23$, $df = 1$, $p = 0.63$, $I^2 = 0.00$, $I^2 = 0.00\%$). An examination of the studentized residuals revealed that none of the studies had a value larger than $±2.69$ and hence there was no indication of outliers in the context of this model. Leave-one-out analyses were not conducted as there were only two studies. According to the Cook’s distances, none of the studies could be considered to be overly influential. With no heterogeneity and only two studies contributing effects, moderator analyses were not appropriate and we were unable to conduct sensitivity analyses by study quality or test for publication bias. There were no dependent effects, so we could not use RVE as a robustness check in this case.
Secondary outcomes

Partial immunisation

We included a total of $k = 9$ studies the analysis. The estimated average outcome based on the random-effects model was $\hat{\mu} = 0.23$ ([95% CI: 0.09 to 0.37], $z = 3.15$, $p = 0.002$), indicating a benefit for the intervention group compared to the control group (see Supplementary Figure 17). A 95% credibility/prediction interval for the true outcomes is given by $-0.20$ to $0.66$. Hence, although the average outcome is estimated to be positive, in some studies the true outcome may in fact be negative.

According to the $Q$-test, the true outcomes appear to be heterogeneous ($Q(8) = 219.11$, $p < 0.001$, $I^2 = 96.351$%). An examination of the studentized residuals revealed that none of the studies had a value larger than $\pm 2.77$ and hence there was no indication of outliers in the context of this model. This was confirmed using a leave-one-out analysis. Likewise, according to the Cook’s distances, none of the studies could be considered to be overly influential. With eight out of the nine studies being assessed as high risk of bias, we were unable to conduct sensitivity analysis by study quality.
Supplementary Figure 17. Forest plot showing the observed outcomes and the estimate of the random-effects model for the impact of community engagement interventions on partial childhood immunisation

As a robustness check, we used robust variance analysis and included all dependent effects in the analysis, totaling 13 effects from the same 9 studies (df = 7.96). The overall average effect was slightly smaller but still significant ($\hat{\mu} = 0.21$ [95% CI: 0.03 to 0.38], $p = .03$). Sensitivity analyses show the effect to be sensitive to all values of Rho.

BCG
We included a total of $k = 12$ studies in the analysis. The estimated average outcome based on the random-effects model was $\hat{\mu} = 0.06$ (95% CI: 0.01 to 0.11), $z = 2.28$, $p = 0.02$, indicating a very small but significant benefit to the intervention participants compared to the control participants (see Supplementary Figure 18). According to the $Q$-test, the true outcomes appear to be heterogeneous ($Q(11) = 84.22$, $p < 0.01$, $I^2 = 86.94\%$). A 95% credibility/prediction interval for the true outcomes is given by $-0.08$ to $0.19$. Hence, although the average outcome is estimated to be positive, in some studies the true outcome may in fact be negative.

An examination of the studentized residuals revealed that one study (Banerjee 2010) had a value larger than $\pm 2.87$ and may be a potential outlier in the context of this model. According to the Cook’s distances, Banerjee (2010) could also be considered to be overly influential. Sensitivity analyses leaving each study out indicated that removing Banerjee (2010) would reduce the overall average effect ($\hat{\mu} = 0.02$ [95% CI: −0.005 to 0.04], and the resulting effect would be non-significant ($z = 1.51$, $p = 0.13$). According to the leave-one-out analysis, there are two additional studies whose removal would result in a non-significant effect; Costa-Font and Parmar (2017) and Rao (2014). Moderator analysis revealed that publication year was a significant predictor such that more recent studies find smaller effects that older studies, with each additional year reducing the size of the effect by .03 standard deviation units ($\hat{\beta} = -0.03$ [95% CI: −0.05 to −0.001], $p = .04$). No other moderators were significant (see Supplementary Table 3). Only three studies were high or moderate quality. When those three studies were synthesised, the resulting effect was reduced ($\hat{\mu} = 0.01$ [95% CI: −0.08 to 0.10]) and no longer significant ($z = 0.22$, $p = 0.82$).
Supplementary Figure 18. Forest plot showing the observed outcomes and the estimate of the random-effects model for the impact of community engagement interventions on BCG vaccination.

A funnel plot of the estimates is shown in Supplementary Figure 18b. The regression test indicated funnel plot asymmetry ($p = 0.04$) but not the rank correlation test ($p = 0.20$).

Supplementary Figure 18b: Funnel plot

As a robustness check, we used robust variance analysis and included all dependent effects in the analysis, totaling 16 effects from the same 12 studies ($df = 7.4$). The overall average effect was slightly smaller and non-significant ($\hat{\mu} = 0.04$ [95% CI: $-0.02$ to $0.10$], $p = .20$). Sensitivity analyses show the effect to be sensitive to all values of Rho.
DPT1

We included a total of $k = 8$ studies in the analysis. The estimated average outcome based on the random-effects model was $\hat{\mu} = 0.04$ ([95% CI: $-0.04$ to $0.11$], $z = 0.99$, $p = 0.32$), indicating no difference between the intervention group and the control group (see Supplementary Figure 19). A 95% credibility/prediction interval for the true outcomes is given by $-0.14$ to $0.21$. Hence, although the average outcome is estimated to be positive, in some studies the true outcome may in fact be negative.

According to the $Q$-test, the true outcomes appear to be heterogeneous ($Q(7) = 30.18, p < 0.01, I^2 = 76.81\%$). An examination of the studentized residuals revealed that none of the studies had a value larger than $\pm 2.73$ and hence there was no indication of outliers in the context of this model. This was confirmed through a leave-one-out analysis. According to the Cook’s distances, none of the studies could be considered to be overly influential. When we removed studies assessed as high risk of bias, two studies remained, and the average effect increased slightly ($\hat{\mu} = 0.04$ ([95% CI: $-0.04$ to $0.11$]), but was still non-significant ($z = 0.99$, $p = 0.32$). Again, with only two studies contributing effects, this must be interpreted with caution.

We tested for potential sources of heterogeneity and found that study design was a significant predictor of the effect such that studies using quasi-experimental designs had smaller effects than RCTs by .15 standard deviation units ($\beta = -0.15$ [95% CI: $-0.29$ to $-0.02$], $p = .03$). We also found that there was a significant difference in the size of effects between programmes implemented by government agencies and those that were not, such that programmes implemented by government agencies (either alone or in tandem with another agency) had larger effects than programmes not implemented by a government agency by .17 standard deviation units ($\beta = 0.17$ [95% CI: $0.01$ to $0.34$], $p = .04$). No other moderators were significant (see Supplementary Table 3).

---

### Table

<table>
<thead>
<tr>
<th>Author (Year)</th>
<th># Participants</th>
<th>SMD [95% CI]</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sankar (2013)</td>
<td>1855</td>
<td>-0.17 [-0.29, -0.05]</td>
</tr>
<tr>
<td>Olayo et al. (2014)</td>
<td>16</td>
<td>-0.09 [-1.02, 0.83]</td>
</tr>
<tr>
<td>Gibson et al. (2017*)</td>
<td>766</td>
<td>0.00 [-0.14, 0.14]</td>
</tr>
<tr>
<td>Costa-Font &amp; Parmar (2017)</td>
<td>64171</td>
<td>0.03 [-0.01, 0.02]</td>
</tr>
<tr>
<td>Webster et al. (2019)</td>
<td>1404</td>
<td>0.04 [-0.07, 0.14]</td>
</tr>
<tr>
<td>Borkum et al. (2015)</td>
<td>493</td>
<td>0.08 [-0.09, 0.26]</td>
</tr>
<tr>
<td>Morris et al. (2004)</td>
<td>2827</td>
<td>0.15 [0.08, 0.23]</td>
</tr>
<tr>
<td>Banajee et al. (2020*)</td>
<td>824</td>
<td>0.16 [0.02, 0.30]</td>
</tr>
</tbody>
</table>

---

Supplementary Figure 19. Forest plot showing the observed outcomes and the estimate of the random-effects model for the impact of community engagement interventions on DPT1 vaccination.
As a robustness check, we used robust variance analysis and included all dependent effects in the analysis, totaling 21 effects from the same 8 studies ($df = 5.24$). The overall average effect was smaller and still non-significant ($\hat{\mu} = 0.01$ [95% CI: −0.06 to 0.09], $p = .66$). Sensitivity analyses show the effect to be sensitive to all values of Rho.

**DPT2**

We included a total of $k = 5$ studies in the analysis. The estimated average outcome based on the random-effects model was $\hat{\mu} = 0.07$ (95% CI: 0.01 to 0.12). Therefore, the average outcome differed significantly from zero ($z = 2.23$, $p = 0.03$), indicating small but significant benefit to the treated group compared to the control group. A 95% credibility/prediction interval for the true outcomes is given by 0.01 to 0.12.

According to the $Q$-test, there was no significant amount of heterogeneity in the true outcomes ($Q(4) = 3.34$, $p = 0.50$, $I^2 = 0.00$, $T^2 = 0.00$%). An examination of the studentized residuals revealed that none of the studies had a value larger than ±2.58 and hence there was no indication of outliers in the context of this model. According to the Cook’s distances, none of the studies could be considered to be overly influential. A leave-one-out analyses indicated that the removal or either Banerjee and colleagues (2020) or Sankar (2013) would results in a non-significant average effect. With no heterogeneity present, we did not examine potential sources of variation. When high risk of bias studies were removed, the resulting average effect increased slightly ($\hat{\mu} = 0.11$ [95% CI: 0.01 to 0.20]) and was still significantly different from zero ($z = 2.12$, $p = 0.03$).

**OPVO**

We included a total of $k = 5$ studies in the analysis. The estimated average outcome based on the random-effects model was $\hat{\mu} = 0.10$ (95% CI: −0.06 to 0.26). Therefore, the average outcome did not differ significantly from zero ($z = 1.24$, $p = 0.22$), indicating no difference between the intervention and control groups (see Supplementary Figure 21). A 95% credibility/prediction interval for the true outcomes is given by −0.25 to 0.45. Hence, although the average outcome is estimated to be positive, in some studies the true outcome may in fact be negative.
According to the Q-test, the true outcomes appear to be heterogeneous ($Q(4) = 45.74, p < 0.01, \hat{\tau}^2 = 0.03, I^2 = 91.26\%$). An examination of the studentized residuals revealed that one study (Sankar 2013) had a value larger than ±2.58 and may be a potential outlier in the context of this model. According to the Cook’s distances, Sankar (2013) could also be considered to be overly influential. Indeed, sensitivity analyses leaving each study out indicated that removing Sankar (2013) would reduce the overall average effect ($\hat{\mu} = 0.01 [95\% CI: -0.01 to 0.03]$) which would become non-significant ($z = 1.02, p = .31$). Leave-one-out analyses confirmed that no other study unduly influenced the outcome.

Exposure to the intervention (in months) was a significant predictor of variation on OPV0 vaccinations such that each additional month of exposure increased the size of the effect by 0.02 standard deviation units ($\beta = .02, [95\% CI: 0.01 to 0.03]; p < .001$). Publication year was also a significant predictor of variation on OPV0 vaccinations such that each additional year decreased the size of the effect by 0.07 standard deviation units ($\beta = -0.07, [95\% CI: -0.11 to -0.02]; p = 0.002$), meaning more recent studies found smaller effects than older studies. With only one study that was not assessed as high risk of bias, we could not conduct a sensitivity analysis by study quality.

Supplementary Figure 21. Forest plot showing the observed outcomes and the estimate of the random-effects model for the impact of community engagement interventions on OPV0 vaccination.

This group of studies was not sufficiently powered for a robustness check using robust variance analysis ($df = 3.76$).

**OPVI**

We included a total of $k = 5$ studies in the analysis. The estimated average outcome based on the random-effects model was $\hat{\mu} = 0.08$ ([95\% CI: 0.004 to 0.15], $z = 2.06, p = 0.04$), indicating a very small but significant benefit to the treated group compared to the control group (see Supplementary Figure 22). A 95\% credibility/prediction interval for the true outcomes is given by $-0.02$ to $0.17$. Hence, although the average outcome is estimated to be positive, in some studies the true outcome may in fact be negative.
According to the $Q$-test, there was no significant amount of heterogeneity in the true outcomes ($Q(4) = 4.77, p = 0.31, \hat{I}^2 = 0.00, I^2 = 16.17\%$), thus we did not test for sources of heterogeneity. An examination of the studentized residuals revealed that none of the studies had a value larger than $\pm 2.58$ and hence there was no indication of outliers in the context of this model. According to the Cook’s distances, none of the studies could be considered to be overly influential. Leave-one-out analyses indicated that the removal of any single study except for Webster et al. (2019) would result in a non-significant average effect.

Supplementary Figure 22. Forest plot showing the observed outcomes and the estimate of the random-effects model for the impact of community engagement interventions on OPV1 vaccination.

This group of studies was not sufficiently powered for a robustness check using robust variance analysis ($df = 3.02$).

OPV2

We included a total of $k = 5$ studies in the analysis. The estimated average outcome based on the random-effects model was $\hat{\mu} = 0.24$ ([95% CI: 0.07 to 0.40], $z = 2.84, p < 0.01$), indicating a small but significant benefit to the treated group compared to the control group (see Supplementary Figure CEOPV2). A 95% credibility/prediction interval for the true outcomes is given by $-0.12$ to $0.59$. Hence, although the average outcome is estimated to be positive, in some studies the true outcome may in fact be negative.

According to the $Q$-test, the true outcomes appear to be heterogeneous ($Q(4) = 22.32, p < 0.01, \hat{I}^2 = 0.03, I^2 = 82.08\%$). An examination of the studentized residuals revealed that none of the studies had a value larger than $\pm 2.58$ and hence there was no indication of outliers in the context of this model. This was confirmed by a leave-one-out analysis. According to the Cook’s distances, none of the studies could be considered to be overly influential. With only one study that was not assessed as high risk of bias, we could not conduct a sensitivity analysis by study quality.

We tested for potential sources of heterogeneity and found several significant moderators in the context of this model (see Supplementary Table 3). Exposure to intervention was significant such that each additional month of exposure increased the average effect by $0.02$ standard deviation units ($\hat{\beta} = 0.02$ [95% CI: 0.01 to 0.02], $p < .0001$). In other words, longer interventions produced larger effects. Publication year was also significant such that older studies reported larger effects than more recent studies. Specifically, each additional year decreased the average effect by $0.05$ standard deviation units ($\hat{\beta} = -0.05$ [95% CI: $-0.08$ to $-0.02$], $p < 0.001$). Finally, region was a significant predictor such that each additional region increased the average effect by $0.03$ standard deviation units ($\hat{\beta} = 0.03$ [95% CI: 0.01 to 0.05], $p < 0.001$).
significant predictor such that Sub-Saharan Africa had smaller effects that South Asia by .26 standard deviation units ($\beta = -0.26$ [95% CI: $-0.39$ to $-0.13$, $p < 0.001$]).

Supplementary Figure 23. Forest plot showing the observed outcomes and the estimate of the random-effects model for the impact of community engagement interventions on OPV2 vaccination.

This group of studies was not sufficiently powered for a robustness check using robust variance analysis ($df = 3.75$).

OPV3

We included total of $k = 9$ studies in the analysis. The estimated average outcome based on the random-effects model was $\bar{\mu} = 0.24$ ([95% CI: 0.09 to 0.40], $z = 3.06$, $p = 0.002$), indicating a moderate and significant benefit to the intervention group compared to the control group (see Supplementary Figure 24). A 95% credibility/prediction interval for the true outcomes is given by $-0.22$ to $0.70$. Hence, although the average outcome is estimated to be positive, in some studies the true outcome may in fact be negative.

According to the $Q$-test, the true outcomes appear to be heterogeneous ($Q(8) = 222.99$, $p < 0.001$, $\tau^2 = 0.05$, $I^2 = 96.41\%$). An examination of the studentized residuals revealed that one study (Sankar 2013) had a value larger than $\pm 2.77$ and may be a potential outlier in the context of this model. According to the Cook’s distances, one study (Sankar 2013) could be considered to be overly influential. Sensitivity analyses leaving each study out indicated that removing Sankar (2013) would substantially reduce the overall average effect ($\bar{\mu} = 0.08$ [95% CI: 0.01 to 0.15]), but the effect would still be positive and significant ($z = 2.39$, $p = .02$). Leave-one-out analyses confirm that no other single study has undue influence on the estimated average effect. When low quality studies were removed, the resulting effect increased, but was no longer significant ($\bar{\mu} = 0.32$ [95% CI: $-0.15$ to $0.79$], $z = 1.32$, $p = 0.19$). With only two studies of high or moderate quality, this result should be interpreted with caution. There were no significant sources of heterogeneity (see Supplementary Table 3).
As a robustness check, we used robust variance analysis and included all dependent effects in the analysis, totaling 10 effects from the same 9 studies (df = 6.74). The overall average effect was identical, but no longer significant ($\hat{\mu} = 0.24$ [95% CI: $-0.02$ to $0.51$], $p = .06$). Sensitivity analyses show the effect to be sensitive to all values of Rho.

**Dropouts**

We included a total of $k = 5$ studies in the analysis. The estimated average outcome based on the random-effects model was $\hat{\mu} = 0.03$ [95% CI: $-0.11$ to $0.16$], $z = 0.36$, $p = 0.72$), indicating no significant difference between the intervention group and the control group (see Supplementary Figure 25). A 95% credibility/prediction interval for the true outcomes is given by $-0.30$ to $0.35$. Hence, although the average outcome is estimated to be positive, in some studies the true outcome may in fact be negative.

According to the $Q$-test, the true outcomes appear to be heterogeneous ($Q(4) = 44.86$, $p < 0.001$, $\hat{\tau}^2 = 0.02$, $I^2 = 91.08\%$). An examination of the studentized residuals revealed that one study (Webster et al. 2019) had a value larger than ±2.58 and may be a potential outlier in the context of this model. Indeed, sensitivity analyses indicated that removing Webster and colleagues (2019) would increase the average effect to $\hat{\mu} = 0.09$ (95% CI: $-0.01$ to $0.19$), but the average outcome still did not differ significantly from zero ($z = 1.70$, $p = 0.09$). Leave-one-out analysis confirmed that no other single study had a disproportionate influence on the estimated average effect. According to the Cook’s distances, none of the studies could be considered to be overly influential. With only one study assessed as having a low risk of bias (Banerjee et al. 2020) and the remaining studies at high risk of bias, we were unable to conduct sensitivity analysis by study quality. Only publication year was a significant predictor such that more recently published studies had higher average effects than older studies. Specifically, each additional year increased the average effect by .33 standard deviation units ($\beta = 0.33$ [95% CI: 0.10 to 0.57], $p = 0.01$).
Supplementary Figure 25. Forest plot showing the observed outcomes and the estimate of the random-effects model for the impact of community engagement interventions on childhood morbidity.

This group of studies was not sufficiently powered for a robustness check using robust variance analysis ($df = 3.98$).

**Morbidity**

We used reports of diarrhea (most typically in the past two weeks) as a proxy for childhood morbidity. In all cases, effects were reverse coded such that positive effects always indicate a benefit to the treated group. Thus, a positive effect here would be interpreted as a reduction in diarrhea among treated participants compared to control participants. A total of $k = 10$ studies were included in the analysis. The estimated average outcome was $\hat{\mu} = 0.01$ ([95% CI: $-0.06$ to $0.08$], $z = 0.34$, $p = 0.73$), indicating no significant difference between intervention group and control group (Supplementary Figure 26). A 95% credibility/prediction interval for the true outcomes is given by $-0.19$ to $0.21$. Hence, although the average outcome is estimated to be positive, in some studies the true outcome may in fact be negative.

According to the $Q$-test, the true outcomes appear to be heterogeneous ($Q(9) = 52.37$, $p < 0.01$, $I^2 = 82.82\%$). When high risk of bias studies were removed, only two studies remained. The resulting estimated average outcome was identical ($\hat{\mu} = 0.01$ [95% CI: $-0.21$ to $0.23$]), and was still non-significant ($z = 0.09$, $p = 0.93$).

An examination of the studentized residuals revealed that one study (Assegaaai et al. 2018) had a value larger than $\pm 2.81$ and may be a potential outlier in the context of this model. According to the Cook’s distances, Assegaaai and colleagues (2018) could also be considered to be overly influential. Indeed, sensitivity analyses leaving each study out indicated that removing Assegaaai and colleagues (2018) would reduce the overall average effect ($\hat{\mu} = -0.02$ [95% CI: $-0.08$ to $0.05$], making the effect negative but still non-significant ($z = -0.47$, $p = .64$). Leave-one-out analyses confirmed that no other single studies had a disproportionate effect on the estimated average effect size. A funnel plot of the estimates is shown in Supplementary Figure 26b. Neither the rank correlation nor the regression test indicated any funnel plot asymmetry ($p = 0.60$ and $p = 0.49$, respectively). We tested for sources of heterogeneity, but there were no significant moderators (see Supplementary Table 3).
Supplementary Figure 26. Forest plot showing the observed outcomes and the estimate of the random-effects model for the impact of community engagement interventions on childhood morbidity.

<table>
<thead>
<tr>
<th>Author (Year)</th>
<th># Participants</th>
<th>SMD  [95% CI]</th>
</tr>
</thead>
<tbody>
<tr>
<td>Admassie et al. (2009)</td>
<td>114</td>
<td>-0.37 [-0.74, -0.01]</td>
</tr>
<tr>
<td>Biemba et al. (2016)</td>
<td>4662</td>
<td>-0.15 [-0.21, -0.09]</td>
</tr>
<tr>
<td>Lee (2015*)</td>
<td>731</td>
<td>-0.10 [-0.25, 0.04]</td>
</tr>
<tr>
<td>Monahan et al. (2020)</td>
<td>5071</td>
<td>-0.01 [-0.06, 0.05]</td>
</tr>
<tr>
<td>Younes et al. (2014)</td>
<td>5246</td>
<td>0.00 [-0.05, 0.06]</td>
</tr>
<tr>
<td>Modi et al. (2019)</td>
<td>3470</td>
<td>0.01 [-0.06, 0.08]</td>
</tr>
<tr>
<td>Webster et al. (2019)</td>
<td>2783</td>
<td>0.06 [-0.02, 0.13]</td>
</tr>
<tr>
<td>Rahman et al. (2008*)</td>
<td>705</td>
<td>0.12 [-0.03, 0.27]</td>
</tr>
<tr>
<td>O’kan et al. (2014)</td>
<td>264</td>
<td>0.14 [-0.10, 0.36]</td>
</tr>
<tr>
<td>Assegaai et al. (2018)</td>
<td>343</td>
<td>0.42 [0.21, 0.63]</td>
</tr>
</tbody>
</table>

RE Model ($Q = 52.37, df = 9, p = 0.00, I^2 = 82.8\%$)

* indicates some concerns or low risk of bias

Supplementary Figure 26b. Funnel plot for studies examining the effect of community engagement interventions on childhood morbidity.

As a robustness check, we used robust variance analysis and included all dependent effects in the analysis, totaling 26 effects from the same 10 studies ($df = 7.69$). The overall average effect was identical and still non-significant ($\hat{\mu} = 0.01$ [95% CI: $-0.09$ to $0.10$], $p = .86$). Sensitivity analyses show the effect to be sensitive to all values of Rho.
Mortality

We included a total of \( k = 6 \) studies in the analysis. The estimated average outcome based on the random-effects model was \( \hat{\mu} = -0.04 \) (95% CI: -0.09 to 0.01; see Supplementary Figure C27). Therefore, the average outcome did not differ significantly from zero \( (z = -1.66, p = 0.10) \), indicating no difference in mortality between the treatment and control groups. A 95% credibility/prediction interval for the true outcomes is given by \(-0.14\) to \(0.06\). Hence, although the average outcome is estimated to be negative, in some studies the true outcome may in fact be positive.

According to the \( Q \)-test, the true outcomes appear to be heterogeneous \( (Q(5) = 19.28, p < 0.01, I^2 = 74.07\%) \). An examination of the studentized residuals revealed that none of the studies had a value larger than \(\pm 2.64\) and hence there was no indication of outliers in the context of this model. According to the Cook’s distances, none of the studies could be considered to be overly influential. Leave-one-out analyses indicated that the removal of Monahan and colleagues (2020) would result in a statistically significant negative effect. When high risk of bias studies were removed, only two studies remained. The resulting estimated average outcome was identical \( (\hat{\mu} = -0.04 [95\% CI: -0.12 to 0.04]) \), and was still non-significant \( (z = -1.07, p = 0.29) \). We tested for sources of heterogeneity, and only publication year was significant such that each additional year increased the effect by .01 standard deviation units \( (\hat{B} = 0.01, p < 0.001 [95\% CI: 0.00 to 0.02]) \). There were no other significant moderators (see Supplementary Table 3).

<table>
<thead>
<tr>
<th>Author (Year)</th>
<th># Participants</th>
<th>SMD [95% CI]</th>
</tr>
</thead>
<tbody>
<tr>
<td>Olken et al. (2014)</td>
<td>264</td>
<td>-0.25 [-0.49, 0.00]</td>
</tr>
<tr>
<td>More et al. (2012*)</td>
<td>15703</td>
<td>-0.08 [-0.12, -0.05]</td>
</tr>
<tr>
<td>Anfeen et al. (2009)</td>
<td>2057</td>
<td>-0.05 [-0.14, 0.03]</td>
</tr>
<tr>
<td>Okeke et al. (2017)</td>
<td>5295</td>
<td>-0.05 [-0.10, 0.00]</td>
</tr>
<tr>
<td>Herrera-Almanza &amp; Rosales-Rueda (2018*)</td>
<td>10772</td>
<td>-0.00 [-0.04, 0.04]</td>
</tr>
<tr>
<td>Monahan et al. (2020)</td>
<td>3763</td>
<td>0.03 [-0.03, 0.10]</td>
</tr>
</tbody>
</table>

RE Model \( (Q = 19.28, df = 5, p = 0.00, I^2 = 74.1\%) \)

* indicates some concerns or low risk of bias

Supplementary Figure 27. Forest plot showing the observed outcomes and the estimate of the random-effects model for the impact of community engagement interventions on childhood morbidity.

As a robustness check, we used robust variance analysis and included all dependent effects in the analysis, totaling 25 effects from the same 6 studies \( (df = 4.07) \). The overall average effect was identical and still non-significant \( (\hat{\mu} = -0.04 [95\% CI: -0.09 to 0.01], p = .12) \). Sensitivity analyses show the effect to be sensitive to all values of Rho.

Immunisation knowledge

We included a total of \( k = 9 \) studies in the analysis. The estimated average outcome based on the random-effects model was \( \hat{\mu} = 0.19 \) (95% CI: 0.07 to 0.31). Therefore, the average outcome differed significantly from zero \( (z = 3.02, p < 0.01) \), indicating a significant benefit to the treatment group compared to the control group (see Supplementary Figure 28). A 95% credibility/prediction interval for the true outcomes is given by \(-0.16\) to \(0.53\). Hence, although the average outcome is estimated to be positive, in some studies the true outcome may in fact be negative.
According to the $Q$-test, the true outcomes appear to be heterogeneous ($Q(8) = 62.87$, $p < 0.01$, $I^2 = 87.28\%$). An examination of the studentized residuals revealed that one study (Banwat et al. 2015) had a value larger than $\pm 2.77$ and may be a potential outlier in the context of this model. Indeed, sensitivity analyses leaving each study out indicated that removing Banwat and colleagues (2015) would reduce the overall average effect ($\hat{\mu} = 0.11$ (95% CI: -0.02 to 0.20), but the effect would still be positive and significant ($z = 2.33, p = .02$). Leave-one-out analyses confirmed no other single study was excessively influential on the average estimated effect. According to the Cook’s distances, none of the studies could be considered to be overly influential.

We examined potential sources of heterogeneity, and study design was a significant predictor of immunisation knowledge such that studies using quasi-experimental designs had larger average effects than RCT designs by .44 standard deviation units ($\beta = 0.44$ [95% CI: 0.23 to 0.64], $p < 0.001$). Region was also a significant predictor, but region was perfectly confounded with study design (e.g. all studies from South Asia were RCTs, while all studies from Sub-Saharan Africa were all quasi-experimental designs), so the effects were larger in Sub-Saharan Africa than from South Asia by the same .44 standard deviation units ($\beta = 0.44$ [95% CI: 0.23 to 0.64], $p < 0.001$). No other moderators were significant (see Supplementary Table 3).

### Table

<table>
<thead>
<tr>
<th>Author (Year)</th>
<th># Participants</th>
<th>SMD [95% CI]</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pramanik (2020)</td>
<td>1450</td>
<td>-0.08 [-0.18, 0.03]</td>
</tr>
<tr>
<td>Banjee et al. (2020*)</td>
<td>1482</td>
<td>-0.01 [-0.12, 0.09]</td>
</tr>
<tr>
<td>Gurley et al. (2020)</td>
<td>1243</td>
<td>0.05 [-0.05, 0.17]</td>
</tr>
<tr>
<td>Andersson et al. (2009)</td>
<td>5564</td>
<td>0.09 [0.04, 0.14]</td>
</tr>
<tr>
<td>Murthy et al. (2019)</td>
<td>505</td>
<td>0.11 [-0.07, 0.26]</td>
</tr>
<tr>
<td>Ocho et al. (2011)</td>
<td>358</td>
<td>0.25 [0.04, 0.46]</td>
</tr>
<tr>
<td>Johri et al. (2020)</td>
<td>159</td>
<td>0.34 [0.03, 0.65]</td>
</tr>
<tr>
<td>Diopolu (2017)</td>
<td>222</td>
<td>0.63 [0.27, 0.79]</td>
</tr>
<tr>
<td>Banwat et al. (2015)</td>
<td>357</td>
<td>0.71 [0.50, 0.92]</td>
</tr>
</tbody>
</table>

RE Model ($Q = 62.87$, $df = 8$, $p = 0.00$, $I^2 = 87.3\%$) $0.19$ [0.07, 0.31]

* indicates some concerns or low risk of bias

Supplementary Figure 28. Forest plot showing the observed outcomes and the estimate of the random-effects model for the impact of community engagement interventions on childhood morbidity.

As a robustness check, we used robust variance analysis and included all dependent effects in the analysis, totaling 13 effects from the same 9 studies ($df = 7.69$). The overall average effect slightly smaller and become non-significant ($\hat{\mu} = 0.17$ [95% CI: -0.02 to 0.37], $p = .07$). Sensitivity analyses show the effect to be sensitive to all values of Rho.

### Immunisation attitudes

We included a total of $k = 6$ studies in the analysis. The estimated average outcome was $\hat{\mu} = 0.14$ (95% CI: $-0.03$ to 0.31), $z = 1.60, p = 0.11$), indicating no significant different between the intervention and control groups (see Supplementary Figure 29). A 95% credibility/prediction interval for the true outcomes is given by $-0.27$ to 0.54.
Hence, although the average outcome is estimated to be positive, in some studies the true outcome may in fact be negative.

According to the $Q$-test, the true outcomes appear to be heterogeneous ($Q(5) = 50.43, p < 0.01, \hat{t}^2 = 0.04, I^2 = 90.09\%$). An examination of the studentized residuals revealed that one study (Banwat et al. 2015) had a value larger than ±2.64 and may be a potential outlier in the context of this model. Indeed, sensitivity analyses leaving each study out indicated that removing Banwat and colleagues (2015) would reduce the overall average effect (μ̂ = 0.04 [95% CI: -0.05 to 0.14]), but the effect would still be positive and non-significant ($z = 0.92, p = .36$). Leave-one-out analyses also indicated that the removal of Dipeolu (2017) would result in a significant positive effect (μ̂ = 0.21 [95% CI: 0.04 to 0.38]). According to the Cook’s distances, none of the studies could be considered to be overly influential. With only one study not assessed as high risk of bias, we were unable to conduct a sensitivity analysis by study quality. We tested for sources of heterogeneity, but of the moderators we were able to test, none were significant (see Supplementary Table 3).

<table>
<thead>
<tr>
<th>Author (Year)</th>
<th># Participants</th>
<th>SMD [95% CI]</th>
</tr>
</thead>
<tbody>
<tr>
<td>Dipeolu (2017)</td>
<td>222</td>
<td>-0.32 [-0.58, -0.05]</td>
</tr>
<tr>
<td>Guney et al. (2020)</td>
<td>1243</td>
<td>0.02 [-0.09, 0.14]</td>
</tr>
<tr>
<td>Andersson et al. (2009)</td>
<td>5636</td>
<td>0.07 [0.02, 0.12]</td>
</tr>
<tr>
<td>Johann et al. (2020)</td>
<td>160</td>
<td>0.10 [-0.21, 0.41]</td>
</tr>
<tr>
<td>Banerjee et al. (2020)</td>
<td>1462</td>
<td>0.14 [0.04, 0.24]</td>
</tr>
<tr>
<td>Banwat et al. (2015)</td>
<td>299</td>
<td>0.02 [0.59, 1.65]</td>
</tr>
</tbody>
</table>

RE Model ($Q = 50.43, df = 5, p = 0.00, I^2 = 90.1\%$) 0.14 [-0.03, 0.31]

* indicates some concerns or low risk of bias

Supplementary Figure 29. Funnel plot for studies examining the effect of community engagement interventions on immunisation attitudes

There were no dependent effects, so we did not complete a robustness check using RVE for this analysis.

**Vaccination card availability**

We included total of $k = 4$ studies in the analysis. The estimated average outcome was $\hat{\mu} = -0.01$ (95% CI: -0.05 to 0.02). The average outcome did not differ significantly from zero ($z = -0.84, p = 0.40$), indicating no difference between the intervention group and the control group on vaccination card availability (see Supplementary Figure 30). A 95% credibility/prediction interval for the true outcomes is given by -0.05 to 0.02.

According to the $Q$-test, there was no significant amount of heterogeneity in the true outcomes ($Q(3) = 2.27, p = 0.52, \hat{t}^2 = 0.00, I^2 = 0.00\%$). One study (Herrera-Almanza & Rosales-Rueda 2018) had a relatively large weight compared to the rest of the studies (i.e., weight ≥ 3/k, so a weight at least 3 times as large as having equal weights across studies). However, this was also the only study assessed as having a low risk of bias. The three other studies in the analysis were assessed as high risk of bias. An examination of the studentized residuals revealed that none of...
the studies had a value larger than ±2.50 and hence there was no indication of outliers in the context of this model. This was confirmed by a leave-one-out analysis. According to the Cook’s distances, none of the studies could be considered to be overly influential. With no heterogeneity present, we did not test for moderation.

<table>
<thead>
<tr>
<th>Author (Year)</th>
<th># Participants</th>
<th>SMD [95% CI]</th>
</tr>
</thead>
<tbody>
<tr>
<td>Olayo et al. (2014)</td>
<td>15</td>
<td>-0.19 [-1.11, 0.74]</td>
</tr>
<tr>
<td>Pramanik (2020)</td>
<td>1450</td>
<td>-0.07 [-0.17, 0.04]</td>
</tr>
<tr>
<td>Herrera-Almanza &amp; Rosales-Rueda (2018*)</td>
<td>12218</td>
<td>-0.01 [-0.05, 0.02]</td>
</tr>
<tr>
<td>Webster et al. (2019)</td>
<td>1376</td>
<td>0.04 [-0.06, 0.15]</td>
</tr>
</tbody>
</table>

* indicates some concerns or low risk of bias

Supplementary Figure 30. Forest plot showing the observed outcomes and the estimate of the random-effects model for the impact of community engagement interventions on vaccination card availability.

There were no dependent effects, so we did not complete a robustness check using RVE for this analysis.

Experience and satisfaction with health services

Only two studies using community engagement interventions reported on experience and satisfaction with health services. The estimated average outcome based on the random-effects model was \( \hat{\mu} = 0.04 \) (95% CI: -0.15 to 0.23). Therefore, the average outcome did not differ significantly from zero (\( z = 0.38, p = 0.70 \), see Supplementary Figure 31). Given the small number of studies, this result should be interpreted with caution. A 95% credibility/prediction interval for the true outcomes is given by -0.13 to 0.31. According to the Q-test, the true outcomes appear to be homogeneous (\( Q(1) = 1.93, p = 0.16, \hat{I}^2 = 0.01, I^2 = 48.22\% \)). With only two studies and no heterogeneity, moderator analyses were not appropriate and tests of publication bias are not valid, and we were not powered for a robustness check using RVE. We also did not complete a leave-one-out analysis in this case.
Formal health worker’s motivation, capacity and performance

Only two studies using community engagement interventions reported on formal health worker’s motivation, capacity and performance. The estimated average outcome based on the random-effects model was $\hat{\mu} = 0.11$ (95% CI: $-0.07$ to $0.29$). Therefore, the average outcome did not differ significantly from zero ($z = 0.38$, $p = 0.70$, see Supplementary Figure 32). Given the small number of studies, this result should be interpreted with caution. A 95% credibility/prediction interval for the true outcomes is given by $-0.07$ to $0.29$. According to the $Q$-test, the true outcomes appear to be homogeneous ($Q(1) = 0.04$, $p = 0.84$, $I^2 = 0.00$, $I^2 = 0.00\%$; With only two studies and no heterogeneity, moderator analyses were not appropriate and tests of publication bias are not valid, and we were not powered for a robustness check using RVE. We also did not complete a leave-one-out analysis in this case.

Supplementary Figure 31. Forest plot showing the observed outcomes and the estimate of the random-effects model for community engagement interventions on experience and satisfaction with health services.
Supplementary Figure 32. Forest plot showing the observed outcomes and the estimate of the random-effects model for community engagement interventions on formal health worker’s motivation, capacity and performance.

<table>
<thead>
<tr>
<th>Author (Year)</th>
<th># Participants</th>
<th>SMD</th>
<th>[95% CI]</th>
</tr>
</thead>
<tbody>
<tr>
<td>Lee (2015*)</td>
<td>307</td>
<td>0.09</td>
<td>[-0.13, 0.31]</td>
</tr>
<tr>
<td>Engineer (2016)</td>
<td>182</td>
<td>0.13</td>
<td>[-0.17, 0.44]</td>
</tr>
</tbody>
</table>

RE Model (Q = 0.04, df = 1, p = 0.84; I² = 0%) 0.11 [-0.07, 0.25]

* indicates some concerns or low risk of bias

Standardized Mean Difference (SMD)
Supplementary Table 2: Summary of quantitative result

<table>
<thead>
<tr>
<th></th>
<th>TOTAL SAMPLE</th>
<th>ENGAGEMENT AS THE INTERVENTION</th>
<th>ENGAGEMENT IN THE DESIGN</th>
<th>ENGAGEMENT IN IMPLEMENTATION AUTONOMY</th>
<th>MULTIPLE ENGAGEMENT TYPES</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>RVE (dependent effects)</td>
<td>RE (independent effects)</td>
<td>g, [95% CI], $I^2$, (k)</td>
<td>g, [95% CI], $I^2$, (k)</td>
<td>g, [95% CI], $I^2$, (k)</td>
</tr>
<tr>
<td>FULL IMMUNISATION</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>0.11**, [0.04, 0.18], 87.27, (28; 53)</td>
<td>0.14**, [0.06, 0.23], 94.46, (28)</td>
<td>0.08**, [0.03, 0.13], 70.00, (12)</td>
<td>010*, [0.02, 0.19], 23.83, (5)</td>
<td>0.23, [-0.001, 0.47], 73.07, (2)</td>
</tr>
<tr>
<td>PARTIAL IMMUNISATION</td>
<td>0.21*, [0.03, 0.38], 95.90, (9; 13)</td>
<td>0.23**, [0.09, 0.37], 96.35, (9)</td>
<td>0.31, [-0.25, 0.87], 98.96, (2)</td>
<td>0.14*, [0.01, 0.27], 0.00, (2)</td>
<td>N/A</td>
</tr>
<tr>
<td>MEASLES</td>
<td>0.06*, [0.01, 0.11], 72.46, (20; 34)</td>
<td>0.07**, [0.03, 0.11], 73.64, (20)</td>
<td>0.10***, [0.05, 0.15], 60.29, (10)</td>
<td>0.11*, [0.02, 0.21], 0.00, (2)</td>
<td>0.03, [-0.09, 0.15], 54.84, (2)</td>
</tr>
<tr>
<td>BCG</td>
<td>0.04, [-0.02, 0.10], 79.17, (12; 16)</td>
<td>0.06*, [0.01, 0.11], 86.94, (12)</td>
<td>0.02***, [0.01, 0.03], 00.00, (4)</td>
<td>0.02, [-0.09, 0.13], 0.00, (2)</td>
<td>0.03, [-0.05, 0.11], 0.00, (2)</td>
</tr>
<tr>
<td>DPT1</td>
<td>0.01, [-0.06, 0.09], 62.55, (8; 21)</td>
<td>0.04, [-0.04, 0.11], 76.81, (8)</td>
<td>0.10, [-0.03, 0.22], 90.21, (3)</td>
<td>0.03, [-0.08, 0.14], 0.00, (2)</td>
<td>N/A</td>
</tr>
<tr>
<td>DPT2</td>
<td>Not powered</td>
<td>0.07*, [0.01, 0.12], 0.00, (5)</td>
<td>N/A</td>
<td>0.05, [-0.06, 0.16], 0.00, (2)</td>
<td>N/A</td>
</tr>
<tr>
<td>DPT3</td>
<td>0.10**, [0.05, 0.15], 76.42, (22; 36)</td>
<td>0.10***, [0.06, 0.14], 76.78, (22)</td>
<td>0.09**, [0.03, 0.15], 73.17, (6)</td>
<td>0.11, [-0.05, 0.28], 80.12, (3)</td>
<td>0.20**, [0.06, 0.34], 87.93, (7)</td>
</tr>
<tr>
<td>OPV0</td>
<td>Not powered</td>
<td>0.10, [-0.06, 0.26], 91.26, (5)</td>
<td>N/A</td>
<td>0.01, [-0.13, 0.14], 0.00, (2)</td>
<td>N/A</td>
</tr>
<tr>
<td>OPV1</td>
<td>Not powered</td>
<td>0.08*, [0.004, 0.15], 16.17, (5)</td>
<td>N/A</td>
<td>0.08, [-0.03, 0.19], 0.00, (2)</td>
<td>N/A</td>
</tr>
<tr>
<td>OPV2</td>
<td>Not powered</td>
<td>0.24**, [0.07, 0.40], 82.08, (5)</td>
<td>N/A</td>
<td>0.23, [-0.09, 0.55], 87.44, (2)</td>
<td>N/A</td>
</tr>
<tr>
<td>OPV3</td>
<td>0.24, [-0.02, 0.51], 96.42, (9; 10)</td>
<td><strong>0.24</strong>, [0.09, 0.40], 96.41, (9)</td>
<td>0.16, [-0.02, 0.34], 83.83, (3)</td>
<td>N/A</td>
<td>0.03, [-0.12, 0.18], 70.38, (2)</td>
</tr>
<tr>
<td>--------------------------</td>
<td>--------------------------------------</td>
<td>-------------------------------------</td>
<td>---------------------------------</td>
<td>-----</td>
<td>---------------------------------</td>
</tr>
<tr>
<td>TIMELINESS (ALL)</td>
<td><strong>0.11</strong>, [0.06, 0.16], 39.92, (11; 54)</td>
<td>----</td>
<td>----</td>
<td>----</td>
<td>----</td>
</tr>
<tr>
<td>TIMELINESS (DTP3)</td>
<td>No dependent effects <strong>0.09</strong>, [0.03, 0.14], 0.00, (7)</td>
<td>N/A</td>
<td><strong>0.12</strong>, [0.03, 0.21], 0.00, (4)</td>
<td>0.04, [-0.06, 0.13], 0.00, (2)</td>
<td>N/A</td>
</tr>
<tr>
<td>TIMELINESS (MEASLES)</td>
<td>No dependent effects <strong>0.23</strong>, [0.14, 0.32], 0.00, (2)</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
</tr>
<tr>
<td>TIMELINESS (COMPLETE IMMUNISATIONS)</td>
<td>No dependent effects <strong>0.15</strong>, [0.07, 0.24], 9.66 (5)</td>
<td>N/A</td>
<td><strong>0.15</strong>, [0.004, 0.29], 0.00, (2)</td>
<td>0.38, [-0.28, 1.03], 72.02, (2)</td>
<td>N/A</td>
</tr>
<tr>
<td>DROPOUTS</td>
<td>Not powered</td>
<td>0.03, [-0.11, 0.16], 91.08, (5)</td>
<td>0.02, [-0.03, 0.06], 0.00, (2)</td>
<td>N/A</td>
<td>N/A</td>
</tr>
<tr>
<td>MORBIDITY</td>
<td>0.01, [-0.09, 0.10], 83.32, (10; 26)</td>
<td>0.01, [-0.06, 0.08], 82.82, (10)</td>
<td>-0.004, [-0.11, 0.10], 88.86, (5)</td>
<td>0.05, [-0.06, 0.15], 45.71, (2)</td>
<td>N/A</td>
</tr>
<tr>
<td>MORTALITY</td>
<td>-0.04, [-0.09, 0.01], 69.91, (6; 25)</td>
<td>-0.04, [-0.09, 0.01], 74.07, (6)</td>
<td>-0.04, [-0.11, 0.04], 80.68, (3)</td>
<td>N/A</td>
<td>N/A</td>
</tr>
<tr>
<td>IMMUNISATION KNOWLEDGE</td>
<td>0.17, [-0.02, 0.37], 87.30, (9; 13)</td>
<td><strong>0.19</strong>, [0.07, 0.31], 87.28, (9)</td>
<td>0.20, [-0.08, 0.49], 93.80, (4)</td>
<td><strong>0.31</strong>, [0.04, 0.58], 72.34, (3)</td>
<td>N/A</td>
</tr>
<tr>
<td>IMMUNISATION ATTITUDES</td>
<td>No dependent effects</td>
<td>0.14, [-0.03, 0.31], 90.09, (6)</td>
<td>0.47, [-0.19, 1.13], 96.36, (2)</td>
<td>-0.11, [-0.53, 0.30], 75.61, (2)</td>
<td>N/A</td>
</tr>
<tr>
<td>VACCINATION CARD AVAILABILITY/RETENTION</td>
<td>No dependent effects</td>
<td>-0.01, [-0.05, 0.02], 0.00, (4)</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
</tr>
<tr>
<td>EXPERIENCE AND SATISFACTION WITH HEALTH SERVICES</td>
<td>No dependent effects</td>
<td>0.04, [-0.015, 0.23], 48.22, (2)</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
</tr>
<tr>
<td>FORMAL HEALTH WORKER CAPACITY AND PERFORMANCE</td>
<td>No powered</td>
<td>0.11, [-0.07, 0.29], 0.00, (2)</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
</tr>
</tbody>
</table>
Supplementary Table 3: Summary of all moderator analyses

<table>
<thead>
<tr>
<th>Moderator Category</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Category 1</td>
<td>Details</td>
</tr>
<tr>
<td>Category 2</td>
<td>Details</td>
</tr>
<tr>
<td>Category 3</td>
<td>Details</td>
</tr>
<tr>
<td>EXP INTMO</td>
<td>EVAL INTMO</td>
</tr>
<tr>
<td>-----------</td>
<td>------------</td>
</tr>
<tr>
<td></td>
<td>ALL EFFECTS RVE (K = 56)</td>
</tr>
<tr>
<td></td>
<td>-0.001</td>
</tr>
<tr>
<td></td>
<td>(0.003)</td>
</tr>
<tr>
<td></td>
<td>FULL IMMUNISATION RVE (K = 28)</td>
</tr>
<tr>
<td></td>
<td>0.0001</td>
</tr>
<tr>
<td></td>
<td>(0.002)</td>
</tr>
<tr>
<td></td>
<td>FULL IMMUNISATION IND RE (K = 28)</td>
</tr>
<tr>
<td></td>
<td>-0.0002</td>
</tr>
<tr>
<td></td>
<td>(0.004)</td>
</tr>
<tr>
<td></td>
<td>FULL IMMUNISATION IND EAI (K = 12)</td>
</tr>
<tr>
<td></td>
<td>-0.003*</td>
</tr>
<tr>
<td></td>
<td>(0.002)</td>
</tr>
<tr>
<td></td>
<td>FULL IMMUNISATION IND EID (K = 5)</td>
</tr>
<tr>
<td></td>
<td>N/A C</td>
</tr>
<tr>
<td></td>
<td>FULL IMMUNISATION IND EII (K = 2)</td>
</tr>
<tr>
<td></td>
<td>NP</td>
</tr>
<tr>
<td></td>
<td>FULL IMMUNISATION IND MET (K = 9)</td>
</tr>
<tr>
<td></td>
<td>0.005</td>
</tr>
<tr>
<td></td>
<td>(0.01)</td>
</tr>
<tr>
<td></td>
<td>PARTIAL IMMUNISATION RVE (K = 9)</td>
</tr>
<tr>
<td></td>
<td>NP</td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>PARTIAL IMMUNISATION IND RE (K = 9)</td>
</tr>
<tr>
<td></td>
<td>-0.003</td>
</tr>
<tr>
<td></td>
<td>(0.006)</td>
</tr>
<tr>
<td></td>
<td>PARTIAL IMMUNISATION IND EAI (K = 2)</td>
</tr>
<tr>
<td></td>
<td>NP</td>
</tr>
<tr>
<td></td>
<td>PARTIAL IMMUNISATION IND EID (K = 2)</td>
</tr>
<tr>
<td></td>
<td>NP</td>
</tr>
<tr>
<td></td>
<td>PARTIAL IMMUNISATION IND EII (K = 1)</td>
</tr>
<tr>
<td></td>
<td>NP</td>
</tr>
<tr>
<td></td>
<td>PARTIAL IMMUNISATION IND MET (K = 4)</td>
</tr>
<tr>
<td></td>
<td>0.01</td>
</tr>
<tr>
<td></td>
<td>(0.02)</td>
</tr>
<tr>
<td></td>
<td>MEASLES RVE (K = 20)</td>
</tr>
<tr>
<td></td>
<td>0.0001</td>
</tr>
<tr>
<td></td>
<td>(0.001)</td>
</tr>
<tr>
<td>Vaccine</td>
<td>Stage</td>
</tr>
<tr>
<td>---------</td>
<td>-------</td>
</tr>
<tr>
<td>MEASLES</td>
<td>IND ME</td>
</tr>
<tr>
<td>MEASLES</td>
<td>IND EAI</td>
</tr>
<tr>
<td>MEASLES</td>
<td>IND EID</td>
</tr>
<tr>
<td>MEASLES</td>
<td>IND RE</td>
</tr>
<tr>
<td>DPT2</td>
<td>IND EAI</td>
</tr>
<tr>
<td>DPT3</td>
<td>IND RE</td>
</tr>
<tr>
<td>DPT2</td>
<td>IND EID</td>
</tr>
<tr>
<td>BCG</td>
<td>IND EID</td>
</tr>
<tr>
<td>BCG</td>
<td>IND EII</td>
</tr>
<tr>
<td>BCG</td>
<td>IND MET</td>
</tr>
<tr>
<td>BCG RVE</td>
<td>K = 12</td>
</tr>
<tr>
<td>BCG IND RE</td>
<td>K = 12</td>
</tr>
<tr>
<td>BCG: IND EAI</td>
<td>K = 4</td>
</tr>
<tr>
<td>BCG: IND EID</td>
<td>K = 2</td>
</tr>
<tr>
<td>BCG: IND EII</td>
<td>K = 2</td>
</tr>
<tr>
<td>BCG: IND MET</td>
<td>K = 4</td>
</tr>
<tr>
<td>DPT1: RVE</td>
<td>K = 8</td>
</tr>
<tr>
<td>DPT1: IND RE</td>
<td>K = 8</td>
</tr>
<tr>
<td>DPT1: IND EAI</td>
<td>K = 3</td>
</tr>
<tr>
<td>DPT1: IND EID</td>
<td>K = 2</td>
</tr>
<tr>
<td>DPT1: IND EII</td>
<td>K = 1</td>
</tr>
<tr>
<td>DPT1: IND MET</td>
<td>K = 2</td>
</tr>
<tr>
<td>DPT2 RVE</td>
<td>K = 5</td>
</tr>
<tr>
<td>DPT2 IND RE</td>
<td>K = 5</td>
</tr>
<tr>
<td>DPT2 IND EAI</td>
<td>K = 0</td>
</tr>
<tr>
<td>DPT2 IND EID</td>
<td>K = 2</td>
</tr>
<tr>
<td>DPT2 IND EII</td>
<td>K = 0</td>
</tr>
<tr>
<td>DPT2 IND MET</td>
<td>K = 3</td>
</tr>
<tr>
<td>DPT3 RVE</td>
<td>K = 22</td>
</tr>
<tr>
<td>DPT3 IND RE</td>
<td>K = 22</td>
</tr>
<tr>
<td>DPT3 IND EAI</td>
<td>K = 6</td>
</tr>
<tr>
<td>DPT3 IND EID</td>
<td>K = 6</td>
</tr>
<tr>
<td>DPT3 IND EII</td>
<td>K = 3</td>
</tr>
<tr>
<td>Supplemental material</td>
<td></td>
</tr>
<tr>
<td>-----------------------</td>
<td></td>
</tr>
<tr>
<td><strong>MEASLES TIMELINESS EAI (K = 0)</strong></td>
<td></td>
</tr>
<tr>
<td><strong>MEASLES TIMELINESS EID (K = 0)</strong></td>
<td></td>
</tr>
<tr>
<td><strong>MEASLES TIMELINESS MET (K = 1)</strong></td>
<td></td>
</tr>
<tr>
<td><strong>DPT3 TIMELINESS MET (K = 1)</strong></td>
<td></td>
</tr>
<tr>
<td><strong>DPT3 TIMELINESS EID (K = 4)</strong></td>
<td></td>
</tr>
<tr>
<td><strong>DPT3 TIMELINESS EAI (K = 0)</strong></td>
<td></td>
</tr>
<tr>
<td><strong>DPT3 TIMELINESS RE (K = 7)</strong></td>
<td></td>
</tr>
<tr>
<td><strong>FI TIMELINESS EAI (K = 1)</strong></td>
<td></td>
</tr>
<tr>
<td><strong>FI TIMELINESS RE (K = 5)</strong></td>
<td></td>
</tr>
<tr>
<td><strong>FI TIMELINESS EID (K = 1)</strong></td>
<td></td>
</tr>
<tr>
<td><strong>FI TIMELINESS MET (K = 1)</strong></td>
<td></td>
</tr>
<tr>
<td><strong>OPV0 IND MET (K = 1)</strong></td>
<td></td>
</tr>
<tr>
<td><strong>OPV0 IND EID (K = 2)</strong></td>
<td></td>
</tr>
<tr>
<td><strong>OPV0 IND EAI (K = 1)</strong></td>
<td></td>
</tr>
<tr>
<td><strong>OPV0 IND EII (K = 1)</strong></td>
<td></td>
</tr>
<tr>
<td><strong>OPV0 IND MET (K = 1)</strong></td>
<td></td>
</tr>
<tr>
<td><strong>OPV1 IND MET (K = 5)</strong></td>
<td></td>
</tr>
<tr>
<td><strong>OPV1 IND EID (K = 1)</strong></td>
<td></td>
</tr>
<tr>
<td><strong>OPV1 IND EAI (K = 0)</strong></td>
<td></td>
</tr>
<tr>
<td><strong>OPV1 IND EII (K = 1)</strong></td>
<td></td>
</tr>
<tr>
<td><strong>OPV2 IND MET (K = 2)</strong></td>
<td></td>
</tr>
<tr>
<td><strong>OPV2 IND EID (K = 2)</strong></td>
<td></td>
</tr>
<tr>
<td><strong>OPV2 IND EAI (K = 0)</strong></td>
<td></td>
</tr>
<tr>
<td><strong>OPV2 IND EII (K = 1)</strong></td>
<td></td>
</tr>
<tr>
<td><strong>OPV3 IND MET (K = 2)</strong></td>
<td></td>
</tr>
<tr>
<td><strong>OPV3 IND EID (K = 2)</strong></td>
<td></td>
</tr>
<tr>
<td><strong>OPV3 IND EAI (K = 1)</strong></td>
<td></td>
</tr>
<tr>
<td><strong>OPV3 IND EII (K = 2)</strong></td>
<td></td>
</tr>
<tr>
<td><strong>OPV3 IND EIII (K = 2)</strong></td>
<td></td>
</tr>
<tr>
<td><strong>OPV3 IND RE (K = 9)</strong></td>
<td></td>
</tr>
<tr>
<td><strong>TIMELINESS RVE (K = 11)</strong></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Indicator</th>
<th>K = 7</th>
<th>K = 11</th>
</tr>
</thead>
<tbody>
<tr>
<td>DPT3 TIMELINESS RE</td>
<td>0.003 (0.006)</td>
<td>N/A</td>
</tr>
<tr>
<td>DPT3 TIMELINESS EAI</td>
<td>-0.01*** (0.004)</td>
<td>N/A</td>
</tr>
<tr>
<td>DPT3 TIMELINESS EID</td>
<td>0.07 (0.16)</td>
<td>N/A</td>
</tr>
<tr>
<td>DPT3 TIMELINESS MET</td>
<td>-0.03* (0.01)</td>
<td>N/A</td>
</tr>
<tr>
<td>MEASLES TIMELINESS RE</td>
<td>-0.10 (0.15)</td>
<td>N/A</td>
</tr>
<tr>
<td>MEASLES TIMELINESS EAI</td>
<td>-0.25 (0.15)</td>
<td>N/A</td>
</tr>
<tr>
<td>MEASLES TIMELINESS EID</td>
<td>0.05 (0.17)</td>
<td>N/A</td>
</tr>
<tr>
<td>MEASLES TIMELINESS MET</td>
<td>-0.04 (0.17)</td>
<td>N/A</td>
</tr>
<tr>
<td>OPV0 IND RVE</td>
<td>N/A</td>
<td>N/A</td>
</tr>
<tr>
<td>OPV0 IND RE</td>
<td>0.02*** (0.004)</td>
<td>N/A</td>
</tr>
<tr>
<td>OPV0 IND EAI</td>
<td>0.16 (0.19)</td>
<td>N/A</td>
</tr>
<tr>
<td>OPV0 IND EID</td>
<td>-0.07** (0.02)</td>
<td>N/A</td>
</tr>
<tr>
<td>OPV0 IND EII</td>
<td>-0.20 (0.18)</td>
<td>N/A</td>
</tr>
<tr>
<td>OPV0 IND MET</td>
<td>0.16 (0.19)</td>
<td>N/A</td>
</tr>
<tr>
<td>OPV1 IND RVE</td>
<td>N/A</td>
<td>N/A</td>
</tr>
<tr>
<td>OPV1 IND RE</td>
<td>0.02*** (0.003)</td>
<td>N/A</td>
</tr>
<tr>
<td>OPV1 IND EAI</td>
<td>0.22 (0.18)</td>
<td>N/A</td>
</tr>
<tr>
<td>OPV1 IND EID</td>
<td>-0.05** (0.02)</td>
<td>N/A</td>
</tr>
<tr>
<td>OPV1 IND EII</td>
<td>-0.26*** (0.07)</td>
<td>N/A</td>
</tr>
<tr>
<td>OPV1 IND MET</td>
<td>0.27 (0.07)</td>
<td>N/A</td>
</tr>
<tr>
<td>OPV2 IND RVE</td>
<td>N/A</td>
<td>N/A</td>
</tr>
<tr>
<td>OPV2 IND RE</td>
<td>0.02*** (0.003)</td>
<td>N/A</td>
</tr>
<tr>
<td>OPV2 IND EAI</td>
<td>-0.03 (0.05)</td>
<td>N/A</td>
</tr>
<tr>
<td>OPV2 IND EID</td>
<td>-0.24 (0.31)</td>
<td>N/A</td>
</tr>
<tr>
<td>OPV2 IND EII</td>
<td>0.03 (0.05)</td>
<td>N/A</td>
</tr>
<tr>
<td>OPV2 IND MET</td>
<td>0.26 (0.30)</td>
<td>N/A</td>
</tr>
<tr>
<td>OPV3 IND RVE</td>
<td>N/A</td>
<td>N/A</td>
</tr>
<tr>
<td>OPV3 IND RE</td>
<td>-0.001 (0.01)</td>
<td>N/A</td>
</tr>
<tr>
<td>OPV3 IND EAI</td>
<td>0.17 (0.23)</td>
<td>N/A</td>
</tr>
<tr>
<td>OPV3 IND EID</td>
<td>-0.03 (0.03)</td>
<td>N/A</td>
</tr>
<tr>
<td>OPV3 IND EII</td>
<td>-0.12 (0.23)</td>
<td>N/A</td>
</tr>
<tr>
<td>OPV3 IND MET</td>
<td>0.04 (0.23)</td>
<td>N/A</td>
</tr>
<tr>
<td>FI TIMELINESS RVE</td>
<td>NP</td>
<td>N/A</td>
</tr>
<tr>
<td>FI TIMELINESS RE</td>
<td>-0.01 (0.01)</td>
<td>N/A</td>
</tr>
<tr>
<td>FI TIMELINESS EAI</td>
<td>0.17 (0.23)</td>
<td>N/A</td>
</tr>
<tr>
<td>FI TIMELINESS EID</td>
<td>-0.03 (0.03)</td>
<td>N/A</td>
</tr>
<tr>
<td>FI TIMELINESS MET</td>
<td>-0.12 (0.23)</td>
<td>N/A</td>
</tr>
</tbody>
</table>

BMJ Publishing Group Limited (BMJ) disclaims all liability and responsibility arising from any reliance placed on this supplemental material which has been supplied by the author(s).

<table>
<thead>
<tr>
<th>Supplemental material</th>
<th>BMJ Publishing Group Limited (BMJ) disclaims all liability and responsibility arising from any reliance placed on this supplemental material which has been supplied by the author(s)</th>
</tr>
</thead>
</table>

<table>
<thead>
<tr>
<th>Variable</th>
<th>K = 0</th>
<th>K = 1</th>
<th>K = 2</th>
<th>K = 3</th>
<th>K = 4</th>
<th>K = 5</th>
<th>K = 6</th>
<th>K = 7</th>
<th>K = 8</th>
<th>K = 9</th>
<th>K = 10</th>
</tr>
</thead>
<tbody>
<tr>
<td>FI TIMELINESS EID</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
</tr>
<tr>
<td>FI TIMELINESS EII</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
</tr>
<tr>
<td>FI TIMELINESS MET</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
</tr>
<tr>
<td>DROPOUT RVE (K = 5)</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
</tr>
<tr>
<td>DROPOUT RE (K = 5)</td>
<td>0.04 (0.04)</td>
<td>0.03 (0.03)</td>
<td>N/A (0.12)</td>
<td>-0.050 (0.183)</td>
<td>N/A (0.13)</td>
<td>-0.23 (0.05)</td>
<td>0.05 (0.18)</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
</tr>
<tr>
<td>DROPOUT EAI (K = 2)</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
</tr>
<tr>
<td>DROPOUT EID (K = 0)</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
</tr>
<tr>
<td>DROPOUT EII (K = 1)</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
</tr>
<tr>
<td>DROPOUT MET (K = 2)</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
</tr>
<tr>
<td>MORBIDITY RVE (K = 10)</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
</tr>
<tr>
<td>MORBIDITY RE (K = 10)</td>
<td>0.003 (0.005)</td>
<td>-0.01 (0.01)</td>
<td>-0.04 (0.11)</td>
<td>0.01 (0.01)</td>
<td>NS (ALL COMPARISONS)</td>
<td>-0.08 (0.10)</td>
<td>0.13 (0.15)</td>
<td>0.08 (0.11)</td>
<td>-0.05 (0.13)</td>
<td>N/A</td>
<td>N/A</td>
</tr>
<tr>
<td>MORBIDITY EAI (K = 5)</td>
<td>0.02** (0.01)</td>
<td>-0.01 (0.02)</td>
<td>N/A (0.05)</td>
<td>0.03 (0.05)</td>
<td>N/A (0.22)</td>
<td>0.12 (0.21)</td>
<td>0.26 (0.22)</td>
<td>0.12 (0.27)</td>
<td>N/A</td>
<td>N/A</td>
<td></td>
</tr>
<tr>
<td>MORBIDITY EID (K = 2)</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
</tr>
<tr>
<td>MORBIDITY EII (K = 1)</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
</tr>
<tr>
<td>MORBIDITY MET (K = 2)</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
</tr>
<tr>
<td>MORTALITY: RVE (K = 6)</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
</tr>
<tr>
<td>MORTALITY RE (6)</td>
<td>-0.002 (0.001)</td>
<td>0.002 (0.002)</td>
<td>0.03 (0.05)</td>
<td>0.01*** (0.003)</td>
<td>NS (ALL COMPARISONS)</td>
<td>N/A (0.18)</td>
<td>0.02 (0.05)</td>
<td>0.02 (0.05)</td>
<td>N/A</td>
<td>N/A</td>
<td></td>
</tr>
<tr>
<td>MORTALITY EAI (3)</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
</tr>
<tr>
<td>MORTALITY EID (0)</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
</tr>
<tr>
<td>MORTALITY EII (0)</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
</tr>
<tr>
<td>MORTALITY MET (3)</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
</tr>
<tr>
<td>KNOWLEDGE RVE (K = 7)</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
</tr>
<tr>
<td>KNOWLEDGE RE (K = 9)</td>
<td>-0.02 (0.01)</td>
<td>-0.01 (0.02)</td>
<td>0.44** (0.11)</td>
<td>-0.01 (0.02)</td>
<td>0.44*** (0.11)</td>
<td>N/A (0.18)</td>
<td>0.05 (0.23)</td>
<td>N/A</td>
<td>-0.01 (0.23)</td>
<td>N/A</td>
<td></td>
</tr>
<tr>
<td>KNOWLEDGE EAI (K = 4)</td>
<td>N/A (0.21)</td>
<td>-0.05 (0.05)</td>
<td>0.52E (0.21)</td>
<td>N/A (0.18)</td>
<td>N/A (0.23)</td>
<td>N/A (0.29)</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
<td></td>
</tr>
<tr>
<td>KNOWLEDGE EID (K = 3)</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
</tr>
<tr>
<td>KNOWLEDGE EII (K = 0)</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
</tr>
<tr>
<td>KNOWLEDGE MET (K = 2)</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
</tr>
<tr>
<td>ATTITUDES RVE (NO DEPENDENT EFFECTS)</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
</tr>
<tr>
<td>ATTITUDES RE (K = 6)</td>
<td>0.01 (0.02)</td>
<td>-0.004 (0.02)</td>
<td>0.17 (0.34)</td>
<td>-0.01 (0.04)</td>
<td>0.17 (0.34)</td>
<td>N/A (0.29)</td>
<td>-0.09 (0.43)</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
</tr>
</tbody>
</table>

128
<table>
<thead>
<tr>
<th></th>
<th>NP</th>
<th>NP</th>
<th>NP</th>
<th>NP</th>
<th>NP</th>
<th>NP</th>
<th>NP</th>
<th>NP</th>
<th>NP</th>
<th>NP</th>
<th>NP</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>ATTITUDES EAI (K = 2)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>ATTITUDES EID (K = 2)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>ATTITUDES EII (K = 0)</strong></td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
</tr>
<tr>
<td><strong>ATTITUDES MET (K = 2)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>CARD AVAIL/ RETENTION RVE</strong></td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
</tr>
<tr>
<td>(NO DEPENDENT EFFECTS)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>CARD AVAIL/ RETENTION RE</strong></td>
<td>N/A'</td>
<td>N/A'</td>
<td>N/A'</td>
<td>N/A'</td>
<td>N/A'</td>
<td>N/A'</td>
<td>N/A'</td>
<td>N/A'</td>
<td>N/A'</td>
<td>N/A'</td>
<td>N/A'</td>
</tr>
<tr>
<td>(K = 4)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>CARD AVAIL/ RETENTION EAI</strong></td>
<td>NP</td>
<td>NP</td>
<td>NP</td>
<td>NP</td>
<td>NP</td>
<td>NP</td>
<td>NP</td>
<td>NP</td>
<td>NP</td>
<td>NP</td>
<td>NP</td>
</tr>
<tr>
<td>(K = 1)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>CARD AVAIL/ RETENTION EID</strong></td>
<td>NP</td>
<td>NP</td>
<td>NP</td>
<td>NP</td>
<td>NP</td>
<td>NP</td>
<td>NP</td>
<td>NP</td>
<td>NP</td>
<td>NP</td>
<td>NP</td>
</tr>
<tr>
<td>(K = 0)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>CARD AVAIL/ RETENTION EII</strong></td>
<td>NP</td>
<td>NP</td>
<td>NP</td>
<td>NP</td>
<td>NP</td>
<td>NP</td>
<td>NP</td>
<td>NP</td>
<td>NP</td>
<td>NP</td>
<td>NP</td>
</tr>
<tr>
<td>(K = 1)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>CARD AVAIL/ RETENTION MET</strong></td>
<td>NP</td>
<td>NP</td>
<td>NP</td>
<td>NP</td>
<td>NP</td>
<td>NP</td>
<td>NP</td>
<td>NP</td>
<td>NP</td>
<td>NP</td>
<td>NP</td>
</tr>
<tr>
<td>(K = 2)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>SATISFACTION W/ HEALTH</strong></td>
<td>NP</td>
<td>NP</td>
<td>NP</td>
<td>NP</td>
<td>NP</td>
<td>NP</td>
<td>NP</td>
<td>NP</td>
<td>NP</td>
<td>NP</td>
<td>NP</td>
</tr>
<tr>
<td>WORKERS (K = 2)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>SATISFACTION W/ HEALTH</strong></td>
<td>NP</td>
<td>NP</td>
<td>NP</td>
<td>NP</td>
<td>NP</td>
<td>NP</td>
<td>NP</td>
<td>NP</td>
<td>NP</td>
<td>NP</td>
<td>NP</td>
</tr>
<tr>
<td>WORKERS (K = 1)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>SATISFACTION W/ HEALTH</strong></td>
<td>NP</td>
<td>NP</td>
<td>NP</td>
<td>NP</td>
<td>NP</td>
<td>NP</td>
<td>NP</td>
<td>NP</td>
<td>NP</td>
<td>NP</td>
<td>NP</td>
</tr>
<tr>
<td>WORKERS EAI (K = 1)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>SATISFACTION W/ HEALTH</strong></td>
<td>NP</td>
<td>NP</td>
<td>NP</td>
<td>NP</td>
<td>NP</td>
<td>NP</td>
<td>NP</td>
<td>NP</td>
<td>NP</td>
<td>NP</td>
<td>NP</td>
</tr>
<tr>
<td>WORKERS EID (K = 1)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>SATISFACTION W/ HEALTH</strong></td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
</tr>
<tr>
<td>WORKERS EII (K = 0)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>SATISFACTION W/ HEALTH</strong></td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
</tr>
<tr>
<td>WORKERS MET (K = 0)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>FORMAL HEALTH WORKERS</strong></td>
<td>NP</td>
<td>NP</td>
<td>NP</td>
<td>NP</td>
<td>NP</td>
<td>NP</td>
<td>NP</td>
<td>NP</td>
<td>NP</td>
<td>NP</td>
<td>NP</td>
</tr>
<tr>
<td>MOTIVATION, CAPACITY, AND</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PERFORMANCE (K = 2)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>FORMAL HEALTH WORKERS</strong></td>
<td>NP</td>
<td>NP</td>
<td>NP</td>
<td>NP</td>
<td>NP</td>
<td>NP</td>
<td>NP</td>
<td>NP</td>
<td>NP</td>
<td>NP</td>
<td>NP</td>
</tr>
<tr>
<td>MOTIVATION, CAPACITY, AND</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PERFORMANCE (K = 1)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>FORMAL HEALTH WORKERS</strong></td>
<td>NP</td>
<td>NP</td>
<td>NP</td>
<td>NP</td>
<td>NP</td>
<td>NP</td>
<td>NP</td>
<td>NP</td>
<td>NP</td>
<td>NP</td>
<td>NP</td>
</tr>
<tr>
<td>MOTIVATION, CAPACITY, AND</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PERFORMANCE EAI (K = 1)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>FORMAL HEALTH WORKERS</strong></td>
<td>NP</td>
<td>NP</td>
<td>NP</td>
<td>NP</td>
<td>NP</td>
<td>NP</td>
<td>NP</td>
<td>NP</td>
<td>NP</td>
<td>NP</td>
<td>NP</td>
</tr>
<tr>
<td>MOTIVATION, CAPACITY, AND</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PERFORMANCE EID (K = 1)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

BMJ Publishing Group Limited (BMJ) disclaims all liability and responsibility arising from any reliance placed on this supplemental material which has been supplied by the author(s).
<table>
<thead>
<tr>
<th></th>
<th>FORMAL HEALTH WORKERS MOTIVATION, CAPACITY, AND PERFORMANCE EII (K = 0)</th>
<th>FORMAL HEALTH WORKERS MOTIVATION, CAPACITY, AND PERFORMANCE MET (K = 0)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>--</td>
<td>--</td>
</tr>
</tbody>
</table>

Note. NS = non-significant, NP = not powered, EAI = engagement as the intervention, EID = engagement in the intervention design, EII = engagement in implementation autonomy, MET = multiple engagement types, RVE = robust variance estimation, RE = random effects model using independent effects.

*Not applicable because all studies were in the same moderator group

b Not applicable because only one study was in a different moderator group

c Not applicable because there was no heterogeneity

d Not applicable because the moderator does not apply to the outcome category

e Not applicable because there was missing data

* p < .05, ** p < .01, *** p < .001
Appendix 10: Quantitative results – Engagement as intervention

Appendix 10 presents supplementary data for analyses presented in the main body of the manuscript for the primary outcomes of community engagement interventions using engagement as the intervention, including funnel plots of publication bias and outlier analyses. It also includes a full presentation of the supplementary outcomes that were precluded from the main body of the manuscript due to word limitations.

Full immunisation

Studies that used engagement as the intervention had a significant effect on full childhood immunisation ($\hat{\mu} = 0.08$ [95% CI: 0.03 to 0.13], $z = 3.02$, $p < 0.01$). A 95% credibility/prediction interval for the true outcomes is given by −0.06 to 0.22. Hence, although the average outcome is estimated to be positive, in some studies the true outcome may in fact be negative.

According to the $Q$-test, the true outcomes appear to be heterogeneous ($Q(11) = 36.67$, $p < 0.01$, $I^2 = 70.00\%$). An examination of the studentized residuals revealed that none of the studies had a value larger than ±2.87 and hence there was no indication of outliers in the context of this model. This was confirmed by a leave-one-out analysis. According to the Cook’s distances, none of the studies could be considered to be overly influential. The forest plot is presented below (Supplementary Figure 33). With only one high or moderate quality study, sensitivity analysis by study quality could not be completed for this body of evidence. A funnel plot of the estimates is shown in Figure B1B. Neither the rank correlation nor the regression test indicated any funnel plot asymmetry ($p = 0.20$ and $p = 0.07$, respectively). Exposure to the intervention (in months) was the only significant source of heterogeneity such that for each additional month of intervention exposure, effects decreased by .003 standard deviation units ($B = −0.003, p = 0.049$ [95% CI: −0.01 to −0.0001]).

<table>
<thead>
<tr>
<th>Author (Year)</th>
<th># Participants</th>
<th>SMD [95% CI]</th>
</tr>
</thead>
<tbody>
<tr>
<td>Rahman et al. (2016)</td>
<td>2700</td>
<td>-0.06 [-0.16, -0.00]</td>
</tr>
<tr>
<td>Pramanik (2020)</td>
<td>971</td>
<td>-0.01 [-0.13, 0.12]</td>
</tr>
<tr>
<td>Findlay et al. (2013)</td>
<td>2305</td>
<td>0.02 [-0.06, 0.10]</td>
</tr>
<tr>
<td>Rao (2014)</td>
<td>34326</td>
<td>0.03 [0.01, 0.05]</td>
</tr>
<tr>
<td>More et al. (2017)</td>
<td>2251</td>
<td>0.06 [-0.02, 0.15]</td>
</tr>
<tr>
<td>Igarashi et al. (2010)</td>
<td>348</td>
<td>0.07 [-0.14, 0.28]</td>
</tr>
<tr>
<td>Biemba et al. (2016)</td>
<td>1508</td>
<td>0.14 [0.04, 0.25]</td>
</tr>
<tr>
<td>Assegai et al. (2016)</td>
<td>403</td>
<td>0.15 [-0.05, 0.34]</td>
</tr>
<tr>
<td>Banwot et al. (2015)</td>
<td>340</td>
<td>0.16 [-0.06, 0.37]</td>
</tr>
<tr>
<td>Younes et al. (2014)</td>
<td>566</td>
<td>0.16 [-0.00, 0.33]</td>
</tr>
<tr>
<td>Monahan et al. (2020)</td>
<td>1073</td>
<td>0.18 [0.06, 0.30]</td>
</tr>
<tr>
<td>Banerjee et al. (2020)</td>
<td>826</td>
<td>0.28 [0.14, 0.42]</td>
</tr>
</tbody>
</table>

RE Model ($Q = 36.67, df = 11, p = 0.00; I^2 = 79.0\%$) $0.08$ [0.03, 0.13]

* Indicates some concerns or low risk of bias

Standardized Mean Difference (SMD)
Supplementary Figure 33. Forest plot showing the observed outcomes and the estimate of the random-effects model for interventions with community engagement as the intervention on full immunisation

Supplementary Figure 33b. Forest plot showing the observed outcomes and the estimate of the random-effects model for interventions with community engagement as the intervention on full immunisation

_DPT3_

Studies that used engagement as the intervention also had a significant effect on DPT3 vaccination (\( \hat{\mu} = 0.09 \) [95% CI: 0.03 to 0.15], \( z = 3.01, p < 0.01 \)). A 95% credibility/prediction interval for the true outcomes is given by \(-0.04\) to \(0.22\). Hence, although the average outcome is estimated to be positive, in some studies the true outcome may in fact be negative.

According to the Q-test, the true outcomes appear to be heterogeneous (\( Q(5) = 18.64, p < 0.01, \hat{\tau}^2 = 0.00, I^2 = 73.17\% \)). An examination of the studentized residuals revealed that none of the studies had a value larger than \(\pm 2.64\) and hence there was no indication of outliers in the context of this model. This was confirmed by a leave-one-out analysis. According to the Cook’s distances, none of the studies could be considered to be overly influential. None of the moderators we tested were significant sources of heterogeneity (see Supplementary Table 3 in Appendix 9).
Supplementary Figure 34. Forest plot showing the observed outcomes and the estimate of the random-effects model for interventions with community engagement as the intervention on DPT3

Measles

Studies that used engagement as the intervention also had a significant effect on measles vaccination ($\hat{\mu} = 0.10$ [95% CI: 0.05 to 0.15], $z = 3.89$, $p < 0.001$). A 95% credibility/prediction interval for the true outcomes is given by $-0.02$ to $0.21$. Hence, although the average outcome is estimated to be positive, in some studies the true outcome may in fact be negative.

According to the $Q$-test, the true outcomes appear to be heterogeneous ($Q(9) = 22.67$, $p = 0.007$, $t^2 = 0.003$, $I^2 = 60.29\%$). An examination of the studentized residuals revealed that none of the studies had a value larger than $\pm 2.81$ and hence there was no indication of outliers in the context of this model. This was confirmed by a leave-one-out analysis. According to the Cook’s distances, none of the studies could be considered to be overly influential.

A funnel plot of the estimates is shown in Figure 35b. The regression test indicated funnel plot asymmetry ($p < 0.01$) but not the rank correlation test ($p = 0.38$). None of the moderators we tested were significant sources of heterogeneity (see Supplementary Table 3 in Appendix 9).
**Supplementary Figure 35. Forest plot showing the observed outcomes and the estimate of the random-effects model for interventions with community engagement as the intervention on measles**

<table>
<thead>
<tr>
<th>Author (Year)</th>
<th># Participants</th>
<th>SMD [95% CI]</th>
</tr>
</thead>
<tbody>
<tr>
<td>Rao (2014)</td>
<td>34335</td>
<td>0.03 [0.01, 0.05]</td>
</tr>
<tr>
<td>Monahan et al. (2020)</td>
<td>1067</td>
<td>0.04 [-0.06, 0.16]</td>
</tr>
<tr>
<td>Morris et al. (2014)</td>
<td>456</td>
<td>0.06 [-0.13, 0.23]</td>
</tr>
<tr>
<td>More et al. (2017)</td>
<td>2251</td>
<td>0.07 [0.01, 0.15]</td>
</tr>
<tr>
<td>Jamison (2011*)</td>
<td>3342</td>
<td>0.08 [0.01, 0.15]</td>
</tr>
<tr>
<td>Aileen et al. (2009)</td>
<td>787</td>
<td>0.09 [-0.05, 0.23]</td>
</tr>
<tr>
<td>Carnell et al. (2014)</td>
<td>1942</td>
<td>0.18 [0.09, 0.27]</td>
</tr>
<tr>
<td>Banerjee et al. (2020*)</td>
<td>824</td>
<td>0.19 [0.06, 0.32]</td>
</tr>
<tr>
<td>Assagasti et al. (2018)</td>
<td>103</td>
<td>0.26 [0.07, 0.46]</td>
</tr>
<tr>
<td>Lee (2015*)</td>
<td>109</td>
<td>0.30 [-0.07, 0.67]</td>
</tr>
</tbody>
</table>

RE Model (Q = 22.57, df = 9, p = 0.01, I² = 60.3%)

*indicates some concern or low risk of bias

Supplementary Figure 35b. Funnel plot showing studies of community engagement as the intervention on measles

**Timeliness of full childhood immunisation**

Only one study (Saggurti 2018) fell into this intervention/outcome category, thus we were unable to perform a statistical synthesis. This quasi-experimental study from India found a small but significant impact of their programme on full childhood immunisation ($g = 0.15$ [95% CI: 0.04 to 0.27]), but like most studies, it was assessed as having a high risk of bias.
Timeliness of DPT3 vaccination

No studies examining timeliness of DPT3 vaccinations used engagement as the intervention.

Timeliness of measles vaccination

No studies examining timeliness of measles vaccinations used engagement as the intervention.

Secondary outcomes

Partial immunisation

Only two studies reporting on partial immunisation used interventions with community engagement as the intervention. The estimated average outcome based on the random-effects model was $\hat{\mu} = 0.31$ ([95% CI: −0.25 to 0.87], $z = 1.10, p = 0.27$), indicating no significant difference between the intervention group and the control group on partial immunisation (see Supplementary Figure 36). Given the small number of studies, this result should be interpreted with caution. A 95% credibility/prediction interval for the true outcomes is given by 0.01 to 0.71. According to the Q-test, the true outcomes appear to be heterogeneous ($Q(1) = 95.93, p < 0.001, \hat{\tau}^2 = 0.16, I^2 = 98.96\%$). With only two studies, moderator and leave-one-out analyses were not appropriate and tests of publication bias are not valid.

<table>
<thead>
<tr>
<th>Author (Year)</th>
<th># Participants</th>
<th>SMD</th>
<th>[95% CI]</th>
</tr>
</thead>
<tbody>
<tr>
<td>Rao (2014)</td>
<td>34336</td>
<td>0.03</td>
<td>[0.01, 0.05]</td>
</tr>
<tr>
<td>Banerjee et al. (2010)</td>
<td>1242</td>
<td>0.50</td>
<td>[0.49, 0.71]</td>
</tr>
</tbody>
</table>

* indicates some concerns or low risk of bias

Supplementary Figure 36. Forest plot showing the observed outcomes and the estimate of the random-effects model for interventions with community engagement as the intervention on partial immunisation.

BCG
We included a total of \( k = 4 \) studies the analysis. The estimated average outcome based on the random-effects model was \( \hat{\mu} = 0.02 \) (95% CI: 0.01 to 0.03). Therefore, the average outcome differed significantly from zero \((z = 3.32, p < 0.01)\), indicating a very small but significant benefit to the treated group compared to the control group (see Supplementary Figure 37). A 95% credibility/prediction interval for the true outcomes is given by 0.01 to 0.03. According to the \( Q \)-test, there was no significant amount of heterogeneity in the true outcomes \((Q(3) = 2.98, p = 0.39, \hat{\tau}^2 = 0.00, I^2 = 0.00\%))\), thus we did not examine potential sources of heterogeneity for this model. An examination of the studentized residuals revealed that none of the studies had a value larger than \( \pm 2.50 \) and hence there was no indication of outliers in the context of this model. This was confirmed by a leave one out analysis. According to the Cook’s distances, none of the studies could be considered to be overly influential.

<table>
<thead>
<tr>
<th>Author (Year)</th>
<th># Participants</th>
<th>SMD [95% CI]</th>
</tr>
</thead>
<tbody>
<tr>
<td>Monahan et al (2020)</td>
<td>1057</td>
<td>-0.04 [-0.16, 0.08]</td>
</tr>
<tr>
<td>Costa-Font &amp; Parmar (2017)</td>
<td>95293</td>
<td>0.01 [0.00, 0.03]</td>
</tr>
<tr>
<td>Rao (2014)</td>
<td>34336</td>
<td>0.03 [0.01, 0.05]</td>
</tr>
<tr>
<td>Lee (2015* )</td>
<td>108</td>
<td>0.16 [0.22, 0.63]</td>
</tr>
<tr>
<td>RE Model (( Q = 2.98, df = 3, p = 0.39, I^2 = 0.00% ))</td>
<td>0.02 [0.01, 0.03]</td>
<td></td>
</tr>
</tbody>
</table>

* indicates some concerns or low risk of bias

Supplementary Figure 37. Forest plot showing the observed outcomes and the estimate of the random-effects model for interventions with community engagement as the intervention on BCG vaccination.

**DPT1**

We included a total of \( k = 3 \) studies in the analysis. The estimated average outcome based on the random-effects model was \( \hat{\mu} = 0.10 \) (95% CI: \(-0.03\) to \(0.22\)], \( z = 1.53 \), \( p = 0.13 \)), indicating no difference between the treatment and control groups (see Supplementary Figure 38). A 95% credibility/prediction interval for the true outcomes is given by \(-0.14\) to \(0.33\). Hence, although the average outcome is estimated to be positive, in some studies the true outcome may in fact be negative.

According to the \( Q \)-test, the true outcomes appear to be heterogeneous \((Q(2) = 20.42, p < 0.01, \hat{\tau}^2 = 0.01, I^2 = 90.21\%))\). An examination of the studentized residuals revealed that one study (Costa-Font & Parmar 2017) had a value larger than \( \pm 2.39 \) and may be a potential outlier in the context of this model. Indeed, sensitivity analyses leaving each study out indicated that removing Costa-Font & Parmar (2017) would increase the overall average effect \( \hat{\mu} = 0.16 \) (95% CI: \(0.09\) to \(0.22\)), and the resulting effect would be positive and significant \((z = 4.72, p < .001)\). Leave-one-out analyses indicated there were no other studies whose removal impacted the average effect estimate. According to the Cook’s distances, none of the studies could be considered to be overly influential. With
only three studies, we were unable to conduct moderator analyses or test for publication bias. Only one study (Banerjee et al. 2020) was not assessed as high risk of bias.

Supplementray Figure 38. Forest plot showing the observed outcomes and the estimate of the random-effects model for interventions with community engagement as the intervention on DPT1 vaccination.

**DPT2**

Only one study (Banerjee et al. 2020) fell into this intervention/outcome category, thus we were unable to perform a statistical synthesis. This cluster RCT from India found a small but significant positive effect of their programme on DPT2 vaccination ($g = 0.15 \ [95\% \text{ CI}: 0.01 \text{ to } 0.29]$), but like most studies, it was assessed as having a high risk of bias.

**OPV0**

Only one study (Costa-Font & Parmar 2017) fell into this intervention/outcome category, thus we were unable to perform a statistical synthesis. This quasi-experimental study from India found a null effect of their programme on OPV0 vaccination ($g = 0.01 \ [95\% \text{ CI}: -0.01 \text{ to } 0.03]$), but like most studies, it was assessed as having a high risk of bias.

**OPV1**

There were no studies reporting on OPV1 vaccination that used interventions with community engagement as the intervention.

**OPV2**
There were no studies reporting on OPV2 vaccination that used interventions with community engagement as the intervention.

**OPV3**

We included a total of $k = 3$ studies in the analysis. The estimated average outcome based on the random-effects model was $\hat{\mu} = 0.16$ ([95% CI: −0.02 to 0.34], $z = 1.78$, $p = 0.07$), indicating no significant difference between the intervention group and the control group on OPV3 vaccination (see Supplementary Figure 39). A 95% credibility/prediction interval for the true outcomes is given by −0.16 to 0.48. Hence, although the average outcome is estimated to be positive, in some studies the true outcome may in fact be negative.

According to the $Q$-test, the true outcomes appear to be heterogeneous ($Q(2) = 12.37$, $p = 0.002$, $I^2 = 83.83\%$). An examination of the studentized residuals revealed that one study (Lee 2015) had a value larger than ±2.39 and may be a potential outlier in the context of this model. Indeed, sensitivity analyses leaving each study out indicated that removing Lee (2015) would reduce the overall average effect ($\hat{\mu} = 0.08$ [95% CI: −0.04 to 0.19]), though in either case the effect is not significantly different from zero ($z = 1.29$, $p = 0.20$). Leave-one-out analyses indicated there were no other studies whose removal impacted the average effect estimate. According to the Cook’s distances, none of the studies could be considered to be overly influential. With only three studies contributing effects, moderator analyses were not appropriate.

<table>
<thead>
<tr>
<th>Author (Year)</th>
<th># Participants</th>
<th>SMD</th>
<th>[95% CI]</th>
</tr>
</thead>
<tbody>
<tr>
<td>Rao (2014)</td>
<td>34336</td>
<td>0.03</td>
<td>[0.01, 0.05]</td>
</tr>
<tr>
<td>Monahan et al. (2020)</td>
<td>1073</td>
<td>0.15</td>
<td>[0.03, 0.27]</td>
</tr>
<tr>
<td>Lee (2015*)</td>
<td>109</td>
<td>0.59</td>
<td>[0.22, 0.97]</td>
</tr>
<tr>
<td>RE Model (Q = 12.37, df = 2, p = 0.001; $I^2 = 83.8%$)</td>
<td></td>
<td>0.16</td>
<td>[-0.02, 0.34]</td>
</tr>
</tbody>
</table>

* indicates some concerns or low risk of bias

Supplementary Figure 39. Forest plot showing the observed outcomes and the estimate of the random-effects model for interventions with community engagement as the intervention on OPV3 vaccination.

**Dropouts**

Only two studies reporting on vaccination dropouts used interventions with community engagement as the intervention. The estimated average outcome based on the random-effects model was $\hat{\mu} = 0.02$ ([95% CI: −0.03 to 0.06], $z = 0.63$, $p = 0.53$), indicating no significant difference between the intervention group and the control group.
on vaccination dropouts (see Supplementary Figure 40). Given the small number of studies, this result should be interpreted with caution. According to the Q-test, the true outcomes appear to be homogeneous (Q(1) = 0.03, p = 0.86, $I^2 = 0.00$). With only two studies and no heterogeneity, moderator and leave-one-out analyses were not appropriate and tests of publication bias are not valid.

<table>
<thead>
<tr>
<th>Author (Year)</th>
<th># Participants</th>
<th>SMD [95% CI]</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pramanik (2020)</td>
<td>1432</td>
<td>0.01 [-0.10, 0.11]</td>
</tr>
<tr>
<td>Banerjee et al. (2020)</td>
<td>5561</td>
<td>0.02 [-0.04, 0.07]</td>
</tr>
</tbody>
</table>

RE Model ($Q = 0.03$, df = 1, $I^2 = 0.00$)

* indicates some concerns or low risk of bias

Supplementary Figure 40. Forest plot showing the observed outcomes and the estimate of the random-effects model for interventions with community engagement as the intervention on vaccination dropouts.

**Morbidity**

We used reports of diarrhea (most typically in the past two weeks) as a proxy for childhood morbidity. In all cases, effects were reverse coded such that positive effects always indicate a benefit to the treated group. Thus, a positive effect here would be interpreted as a reduction in diarrhea among treated participants compared to control participants. We included a total of $k = 5$ studies in the analysis. The estimated average outcome based on the random-effects model was $\hat{\mu} = -0.004$ (95% CI: -0.11 to 0.10). Therefore, the average outcome did not differ significantly from zero ($z = -0.08$, $p = 0.94$), indicating no difference between the treatment and control groups on childhood morbidity (see Supplementary Figure 41). A 95% credibility/prediction interval for the true outcomes is given by -0.24 to 0.23. Hence, although the average outcome is estimated to be negative, in some studies the true outcome may in fact be positive.

According to the Q-test, the true outcomes appear to be heterogeneous ($Q(4) = 35.91$, $p < 0.01$, $I^2 = 0.01$, $I^2 = 88.86$%). An examination of the studentized residuals revealed that one study (Assegai et al. 2018) had a value larger than ±2.58 and may be a potential outlier in the context of this model. Indeed, sensitivity analyses leaving each study out indicated that removing Assegai and colleagues (2018) would reduce the overall average effect ($\hat{\mu} = -0.06$ (95% CI: -0.14 to 0.02). While the resulting effect would be negative, the effect would still be non-significant ($z = -1.50$, $p = .13$). Leave-one-out analyses confirmed that no other studies exerted a disproportionate influence on the estimated average effect. According to the Cook’s distances, none of the studies could be considered to be overly influential. There was only one study not assessed as high risk of bias, so we were unable to do a sensitivity analysis by study quality. Exposure to the intervention was a significant source of heterogeneity such that each additional
month of exposure increased the positive impact on morbidity by 0.02 standard deviation units $\hat{\mu} = 0.02$ (95% CI: 0.01 to 0.03, $p = 0.002$). No other moderators were significant (see Supplementary Table 3).

<table>
<thead>
<tr>
<th>Author/Year</th>
<th># Participants</th>
<th>SMD  [95% CI]</th>
</tr>
</thead>
<tbody>
<tr>
<td>Biemba et al. (2016)</td>
<td>4662</td>
<td>-0.15 [-0.21, -0.09]</td>
</tr>
<tr>
<td>Lee (2015*)</td>
<td>731</td>
<td>-0.10 [0.25, 0.04]</td>
</tr>
<tr>
<td>Monahan et al. (2020)</td>
<td>5071</td>
<td>-0.01 [-0.06, 0.05]</td>
</tr>
<tr>
<td>Younes et al (2014)</td>
<td>5246</td>
<td>-0.00 [0.06, 0.05]</td>
</tr>
<tr>
<td>Assegasi et al. (2018)</td>
<td>343</td>
<td>0.42 [0.21, 0.63]</td>
</tr>
</tbody>
</table>

Supplementary Figure 41. Forest plot showing the observed outcomes and the estimate of the random-effects model for interventions with community engagement as the intervention on morbidity.

**Mortality**

We included a total of $k = 3$ studies in the analysis. The estimated average outcome based on the random-effects model was $\hat{\mu} = -0.04$ (95% CI: -0.11 to 0.04). Therefore, the average outcome did not differ significantly from zero ($z = -0.94$, $p = 0.34$), indicating no difference between the treatment and control group (see Supplementary Figure 42). A 95% credibility/prediction interval for the true outcomes is given by $-0.18$ to $0.10$. Hence, although the average outcome is estimated to be negative, in some studies the true outcome may in fact be positive.

According to the $Q$-test, the true outcomes appear to be heterogeneous ($Q(2) = 10.35$, $p < 0.01$, $I^2 = 80.68\%$). An examination of the studentized residuals revealed that one study (Monahan et al. 2020) had a value larger than $\pm 2.39$ and may be a potential outlier in the context of this model. Indeed, sensitivity analyses leaving each study out indicated that removing Monahan and colleagues (2020) would reduce the overall average effect ($\hat{\mu} = 0.01$ [95% CI: -0.14 to 0.16]), but the effect would still be positive and non-significant ($z = 0.10$, $p = .92$). Leave-one-out analyses confirmed that no other studies exerted a disproportionate influence on the estimated average effect. According to the Cook’s distances, none of the studies could be considered to be overly influential. With only three studies contributing effects, we were unable to test for potential sources of heterogeneity.
**Immunisation knowledge**

We included a total of \( k = 4 \) studies in the analysis. The estimated average outcome based on the random-effects model was \( \hat{\mu} = 0.20 \text{ (95\% CI: } -0.08 \text{ to } 0.49) \). Therefore, the average outcome did not differ significantly from zero \((z = 1.41, p = 0.16)\), indicating no difference between the treatment and control groups. A 95\% credibility/prediction interval for the true outcomes is given by \(-0.41\) to \(0.81\). Hence, although the average outcome is estimated to be positive, in some studies the true outcome may in fact be negative.

According to the \( Q \)-test, the true outcomes appear to be heterogeneous \((Q(3) = 48.39, p < 0.01, \tau^2 = 0.08, I^2 = 93.80\%\)). An examination of the studentized residuals revealed that one study (Banwat et al. 2015) had a value larger than \( \pm 2.50 \) and may be a potential outlier in the context of this model. Indeed, sensitivity analyses leaving each study out indicated that removing Banwat and colleagues (2015) would reduce the overall average effect \((\hat{\mu} = 0.02 \text{ (95\% CI: } -0.12 \text{ to } 0.17)\), but the effect would still be positive and non-significant \((z = 0.35, p = .73)\). Leave-one-out analyses confirmed that no other studies exerted a disproportionate influence on the estimated average effect. According to the Cook’s distances, none of the studies could be considered to be overly influential.

We examined potential sources of heterogeneity, and study design was a significant predictor of immunisation knowledge such that studies using quasi-experimental designs had larger average effects than RCT designs by \( .52 \) standard deviation units \((\hat{\beta} = 0.52 \text{ [95\% CI: } 0.11 \text{ to } 0.94], p = 0.01)\). Region was also a significant predictor, but region was perfectly confounded with study design (e.g. all studies from South Asia were RCTs, while all studies from Sub-Saharan Africa were all quasi-experimental designs), so the effects were larger in Sub-Saharan Africa than from South Asia by the same \( .52 \) standard deviation units \((\beta = 0.52 \text{ [95\% CI: } 0.11 \text{ to } 0.94], p = 0.01)\). No other moderators were significant (see Supplementary Table 3 in Appendix 9).
Immunisation attitudes

Only two studies reporting on immunisation attitudes used interventions with community engagement as the intervention. The estimated average outcome based on the random-effects model was $\hat{\mu} = 0.47$ ([95% CI: −0.19 to 1.13], $z = 1.40$, $p = 0.16$), indicating no significant difference between the intervention group and the control group (see Supplementary Figure 44). Given the small number of studies, this result should be interpreted with caution. According to the Q-test, the true outcomes appear to be heterogeneous ($Q(1) = 27.51$, $p < 0.001$, $\hat{\tau}^2 = 0.22$, $I^2 = 96.36\%$; With only two studies and no heterogeneity, moderator and leave-one-out analyses were not appropriate and tests of publication bias are not valid.)
Supplementary Figure 44. Forest plot showing the observed outcomes and the estimate of the random-effects model for interventions with community engagement as the intervention on immunisation attitudes.

Vaccination card retention

Only one study (Pramanik 2020) fell into this intervention/outcome category, thus we were unable to perform a statistical synthesis. This cluster RCT from India found a null effect of their programme on vaccination card retention/availability ($g = -0.07$ [95% CI: -0.17 to 0.04]), but like most studies, it was assessed as having a high risk of bias.

Experience and satisfaction with health services

Only one study (Mohanan et al. 2020) fell into this intervention/outcome category, thus we were unable to perform a statistical synthesis. This cluster RCT from India did not find a significant effect of their programme on satisfaction with health services ($g = -0.02$ [95% CI: -0.08 to 0.04]), but like most studies, it was assessed as having a high risk of bias.

Formal health worker’s motivation, capacity and performance

Only one study (Lee 2015) fell into this intervention/outcome category, thus we were unable to perform a statistical synthesis. This cluster RCT from Zambia did not find a significant effect of their programme on formal health worker’s motivation, capacity, and performance ($g = 0.09$ [95% CI: -0.13 to 0.31]). This study was assessed as having some concerns related to risk of bias.
Appendix 11: Quantitative results - community engagement in intervention design

Appendix 11 presents supplementary data for analyses presented in the main body of the manuscript for the primary outcomes of community engagement interventions using engagement in the intervention design, including funnel plots of publication bias and outlier analyses. It also includes a full presentation of the supplementary outcomes that were precluded from the main body of the manuscript due to word limitations.

Full immunisation

When studies used community engagement in the design, there was a significant effect on full childhood immunisation ($\hat{\mu} = 0.10$ [95% CI: 0.02 to 0.19], $z = 2.40$, $p = 0.02$). A 95% credibility/prediction interval for the true outcomes is given by −0.02 to 0.23. Hence, although the average outcome is estimated to be positive, in some studies the true outcome may in fact be negative.

According to the $Q$-test, there was no significant amount of heterogeneity in the true outcomes ($Q(4) = 5.25$, $p = 0.26$, $I^2 = 0.00$, $I^2 = 23.83\%$). An examination of the studentized residuals revealed that none of the studies had a value larger than ±2.58 and hence there was no indication of outliers in the context of this model. Leave-one-out analyses indicated that removal of Murthy et al. (2019), Borkum et al. (2015) or Rahman et al. (2008) would result in a non-significant effect estimate. According to the Cook’s distances, none of the studies could be considered to be overly influential.

With no heterogeneity among effects, moderator analyses were not appropriate. When low quality studies were removed, two studies remained, and the summary effect increased ($\hat{\mu} = 0.14$ [95% CI: −0.01 to 0.29], but the effect was no longer significant, ($z = 1.89$, $p = 0.06$).

Supplementary Figure 45. Forest plot showing the observed outcomes and the estimate of the random-effects model for interventions with community engagement in the design on full immunisation.
DPT3

When studies used community engagement in the design, there was not a significant effect on DPT3 vaccination ($\hat{\mu} = 0.04$ [95% CI: $-0.01$ to $0.08$], $z = 1.69, p = 0.09$). According to the $Q$-test, there was no significant amount of heterogeneity in the true outcomes ($Q(5) = 4.36, p = 0.50, \hat{I}^2 = 0.00, I^2 = 0.00\%$). An examination of the studentized residuals revealed that none of the studies had a value larger than $\pm 2.64$ and hence there was no indication of outliers in the context of this model. According to the Cook’s distances, none of the studies could be considered to be overly influential. With no heterogeneity, we did not perform moderator or leave-one-out analyses, or compute a prediction interval. With only one study of low or medium quality, we were also unable to complete sensitivity analyses for this body of evidence.

<table>
<thead>
<tr>
<th>Author (Year)</th>
<th># Participants</th>
<th>SMD</th>
<th>[95% CI]</th>
</tr>
</thead>
<tbody>
<tr>
<td>Engineer (2016)</td>
<td>143</td>
<td>-0.14</td>
<td>[-0.47, 0.19]</td>
</tr>
<tr>
<td>Carmichael et al. (2019)</td>
<td>1558</td>
<td>0.00</td>
<td>[-0.10, 0.10]</td>
</tr>
<tr>
<td>Modi et al. (2019)</td>
<td>3470</td>
<td>0.02</td>
<td>[-0.05, 0.08]</td>
</tr>
<tr>
<td>Siddiqi (2020)</td>
<td>963</td>
<td>0.02</td>
<td>[-0.11, 0.15]</td>
</tr>
<tr>
<td>Gibson et al. (2017)</td>
<td>766</td>
<td>0.04</td>
<td>[-0.10, 0.18]</td>
</tr>
<tr>
<td>Nagar et al. (2020)</td>
<td>2243</td>
<td>0.10</td>
<td>[0.02, 0.18]</td>
</tr>
</tbody>
</table>

RE Model ($Q = 4.36, df = 5, p = 0.50; \hat{I}^2 = 0.0\%$) 0.04 [-0.01, 0.00]

* indicates some concerns or low risk of bias

Supplementary Figure 46. Forest plot showing the observed outcomes and the estimate of the random-effects model for interventions with community engagement in the design on DPT3 vaccination

Measles

When studies used community engagement in the design, there was a significant effect on measles vaccination ($\hat{\mu} = 0.11$ [95% CI: $0.02$ to $0.21$], $z = 2.36, p = 0.02$). According to the Q-test, there was no significant amount of heterogeneity in the true outcomes ($Q(1) = 0.56, p = 0.46, \hat{I}^2 = 0.00, I^2 = 0.00\%$). With only two studies and no heterogeneity among effects, moderator and leave-one-out analyses were not appropriate.
Supplementary Figure 47. Forest plot showing the observed outcomes and the estimate of the random-effects model for interventions with community engagement in the design on measles vaccination

**Timeliness of full childhood immunisation**

When studies used community engagement in the design, there was also a significant effect on timeliness full childhood immunisation \( \hat{\mu} = 0.15 \) (95% CI: 0.004 to 0.29), \( z = 2.01, p = 0.04 \). According to the Q-test, there was no significant amount of heterogeneity in the true outcomes (Q(1) = 0.81, \( \hat{\tau}^2 = 0.00 \), I\(^2\) = 0.00%). Both studies were assessed as high risk of bias. With only two studies and no heterogeneity, moderator and leave-one-out analyses were not appropriate.

---

### Author (Year) # Participants SMD [95% CI]

<table>
<thead>
<tr>
<th>Author (Year)</th>
<th># Participants</th>
<th>SMD [95% CI]</th>
</tr>
</thead>
<tbody>
<tr>
<td>Siddique (2020)</td>
<td>963</td>
<td>0.08 [-0.04, 0.21]</td>
</tr>
<tr>
<td>Gibson et al. (2017)</td>
<td>766</td>
<td>0.15 [0.01, 0.30]</td>
</tr>
<tr>
<td>RE Model (Q = 0.55, df = 1, ( \hat{\tau}^2 = 0.00 ))</td>
<td></td>
<td>0.11 [0.02, 0.21]</td>
</tr>
</tbody>
</table>

* indicates some concerns or low risk of bias

---

Standardized Mean Difference (SMD)
Supplementary Figure 48. Forest plot showing the observed outcomes and the estimate of the random-effects model for interventions with community engagement in the design on the timeliness of full childhood immunisation

**Timeliness of DPT3 vaccination**

When studies used community engagement in the design, there was also a significant effect on the timeliness of DPT3 vaccinations ($\bar{\mu} = 0.12 \ [95\% \ CI: 0.03\ to\ 0.21]$, $z = 2.73$, $p < 0.01$). According to the Q-test, there was no significant amount of heterogeneity in the true outcomes ($Q(3) = 2.62$, $p = 0.45$, $\tau^2 = 0.00$, $I^2 = 0.00\%$). An examination of the studentized residuals revealed that none of the studies had a value larger than ±2.50 and hence there was no indication of outliers in the context of this model. A leave-one-out analysis indicated that removing Borkum et al. (2014) or Gibson et al. (2017) would result in a non-significant average estimated effect. According to the Cook’s distances, none of the studies could be considered to be overly influential. With no heterogeneity indicated, we did not test for moderation.

**Supplementary Table 48.**

<table>
<thead>
<tr>
<th>Author (Year)</th>
<th># Participants</th>
<th>SMD [95% CI]</th>
</tr>
</thead>
<tbody>
<tr>
<td>Domok et al. (2019*)</td>
<td>662</td>
<td>0.02 [-0.13, 0.18]</td>
</tr>
<tr>
<td>Nagar et al. (2010)</td>
<td>99</td>
<td>0.10 [-0.29, 0.49]</td>
</tr>
<tr>
<td>Borkum et al. (2014)</td>
<td>468</td>
<td>0.16 [-0.02, 0.34]</td>
</tr>
<tr>
<td>Gibson et al. (2017*)</td>
<td>766</td>
<td>0.19 [0.05, 0.33]</td>
</tr>
</tbody>
</table>

(RE Model ($Q = 2.62$, df = 3, $p = 0.45$, $I^2 = 0.0\%$))

* indicates some concerns or low risk of bias

Supplementary Figure 49. Forest plot showing the observed outcomes and the estimate of the random-effects model for interventions with community engagement in the design of the intervention on the timeliness of DPT3 vaccinations.

**Timeliness of measles vaccination**

Only one study reporting on timeliness of measles vaccination used an intervention with engagement in the design (Gibson et al. 2017). This RCT conducted in Kenya found a small but significant positive effect ($g = 0.20 \ [95\% \ CI: 0.06\ to\ 0.34]$). This study was assessed as having some concerns of bias.

**Secondary outcomes**

**Partial immunisation**
Only two studies reporting on partial immunisation used interventions with community engagement in the intervention design. The estimated average outcome based on the random-effects model was $\hat{\mu} = 0.14$ ([95% CI: 0.01 to 0.27], $z = 2.05$, $p = 0.04$), indicating a small but significant benefit to the intervention group compared to the control group (see Supplementary Figure 50). Given the small number of studies, this result should be interpreted with caution. According to the Q-test, the true outcomes appear to be homogeneous ($Q(1) = 0.91$, $p = 0.34$, $\tau^2 = 0.00$, $I^2 = 0.00\%$). With only two studies and no heterogeneity, moderator and leave-one-out analyses were not appropriate and tests of publication bias are not valid.

![Forest plot showing the observed outcomes and the estimate of the random-effects model for interventions with community engagement in the design of the intervention on partial childhood immunisation.](image)

**BCG**

Only two studies reporting on BCG vaccination used interventions with community engagement in the intervention design. The estimated average outcome based on the random-effects model was $\hat{\mu} = 0.02$ ([95% CI: $-0.09$ to $0.13$], $z = 0.32$, $p = 0.75$), indicating no difference between the intervention group and the control group (see Supplementary Figure 51). Given the small number of studies, this result should be interpreted with caution. According to the Q-test, the true outcomes appear to be homogeneous ($Q(1) = 0.11$, $p = 0.75$, $\tau^2 = 0.00$, $I^2 = 0.00\%$). With only two studies and no heterogeneity, moderator and leave-one-out analyses were not appropriate and tests of publication bias are not valid.
Only two studies reporting on DPT1 vaccination used interventions with community engagement in the intervention design. The estimated average outcome based on the random-effects model was $\hat{\mu} = 0.03$ (95% CI: $-0.08$ to $0.14$). Therefore, the average outcome did not differ significantly from zero ($z = 0.60, p = 0.55$, see Supplementary Figure 52), indicating no difference between the treatment group and the control group on DPT1 vaccination. Given the small number of studies, this result should be interpreted with caution. According to the Q-test, the true outcomes appear to be homogeneous ($Q(1) = 0.48, p = 0.49, I^2 = 0.00, l^2 = 0.00\%$). With only two studies and no heterogeneity, moderator and leave-one-out analyses were not appropriate and tests of publication bias are not valid.
Supplementary Figure 52. Forest plot showing the observed outcomes and the estimate of the random-effects model for interventions with community engagement in the design on the intervention on DPT1 vaccination.

DPT2

Only two studies reporting on DPT2 vaccination used interventions with community engagement in the intervention design. The estimated average outcome based on the random-effects model was $\hat{\mu} = 0.05$ (95% CI: $-0.06$ to $0.16$). Therefore, the average outcome did not differ significantly from zero ($z = 0.83$, $p = 0.41$, see Supplementary Figure 53), indicating no difference between the treatment group and the control group on DPT2 vaccination. Given the small number of studies, this result should be interpreted with caution. According to the Q-test, the true outcomes appear to be homogeneous ($Q(1) = 0.07$, $p = 0.41$, $I^2 = 0.00$, $I^2 = 0.00\%$). With only two studies and no heterogeneity, moderator and leave-one-out analyses were not appropriate and tests of publication bias are not valid.
OPV0

Only two studies reporting on immunisation attitudes used interventions with community engagement in the intervention design. The estimated average outcome based on the random-effects model was $\hat{\mu} = 0.01$ (95% CI: $-0.13$ to $0.14$). Therefore, the average outcome did not differ significantly from zero ($z = 0.08, p = 0.94$, see Supplementary Figure 54). Given the small number of studies, this result should be interpreted with caution. According to the Q-test, the true outcomes appear to be heterogeneous ($Q(1) = 0.01, p = 0.93, I^2 = 0.00, I^2 = 0.00\%$). With only two studies and no heterogeneity, moderator and leave-one-out analyses were not appropriate and tests of publication bias are not valid.

---

**Supplementary Figure 53. Forest plot showing the observed outcomes and the estimate of the random-effects model for interventions with community engagement in the design on the intervention on DPT2 vaccination.**

<table>
<thead>
<tr>
<th>Author (Year)</th>
<th># Participants</th>
<th>SMD [95% CI]</th>
</tr>
</thead>
<tbody>
<tr>
<td>Borium et al. (2015)</td>
<td>493</td>
<td>0.03 [-0.15, 0.20]</td>
</tr>
<tr>
<td>Gibson et al. (2017)</td>
<td>768</td>
<td>0.06 [-0.06, 0.20]</td>
</tr>
</tbody>
</table>

*indicates some concerns or low risk of bias

**RE Model** ($Q = 0.07, df = 1, p = 0.80; I^2 = 0.0\%$)

$0.05 [-0.06, 0.16]$
Supplementary Figure 54. Forest plot showing the observed outcomes and the estimate of the random-effects model for interventions with community engagement in the design on the intervention on OPV0 vaccination.

**OPV1**

Only two studies reporting on OPV1 vaccination used interventions with community engagement in the intervention design. The estimated average outcome based on the random-effects model was $\hat{\mu} = 0.08$ (95% CI: $-0.03$ to $0.19$). Therefore, the average outcome did not differ significantly from zero ($z = 1.44, p = 0.15$, see Supplementary Figure 55). Given the small number of studies, this result should be interpreted with caution. According to the Q-test, the true outcomes appear to be homogeneous ($Q(1) = 0.20, p = 0.65$, $\hat{\xi}^2 = 0.00$, $I^2 = 0.00\%$). With only two studies, moderator and leave-one-out analyses were not appropriate and tests of publication bias are not valid.
Supplementary Figure 55. Forest plot showing the observed outcomes and the estimate of the random-effects model for interventions with community engagement in the design on the intervention on OPV1 vaccination.

**OPV2**

Only two studies reporting on OPV2 vaccination used interventions with community engagement in the intervention design. The estimated average outcome based on the random-effects model was $\hat{\mu} = 0.23$ (95% CI: −0.09 to 0.55). Therefore, the average outcome did not differ significantly from zero ($z = 1.43, p = 0.15$, see Supplementary Figure 56), indicating no significant difference between the intervention group and the control group on OPV2 vaccinations. Given the small number of studies, this result should be interpreted with caution. According to the Q-test, the true outcomes appear to be heterogeneous ($Q(1) = 7.96, p = 0.004, \hat{\tau}^2 = 0.05, I^2 = 87.44\%$). With only two studies, moderator and leave-one-out analyses were not appropriate and tests of publication bias are not valid.
Supplementary Figure 56. Forest plot showing the observed outcomes and the estimate of the random-effects model for interventions with community engagement in the design on the intervention on OPV2 vaccination.

**OPV3**

Only one study (Borkum 2015) fell into this intervention/outcome category, thus we were unable to perform a statistical synthesis. This cluster RCT from India found a small but significant positive effect of their programme on OPV3 vaccination ($g = 0.19 \ [95\% \ CI: 0.01 \ to \ 0.36]$), but again, it was assessed as having a high risk of bias.

**Dropouts**

There were no studies using engagement in the intervention design that reported on vaccination dropouts.

**Morbidity**

We used reports of diarrhea (most typically in the past two weeks) as a proxy for childhood morbidity. In all cases, effects were reverse coded such that positive effects always indicate a benefit to the treated group. Thus, a positive effect here would be interpreted as a reduction in diarrhea among treated participants compared to control participants. Only two studies reporting on childhood morbidity used interventions with community engagement in the intervention design. The estimated average outcome based on the random-effects model was $\hat{\mu} = 0.05$ (95% CI: $-0.06$ to $0.15$). Therefore, the average outcome did not differ significantly from zero ($z = 0.88, p = 0.38$, see Supplementary Figure 57). Given the small number of studies, this result should be interpreted with caution. According to the Q-test, the true outcomes appear to be homogeneous ($Q(1) = 1.84, p = 0.17, I^2 = 0.003, I^2 = 45.71\%$). With only two studies, moderator and leave-one-out analyses were not appropriate and tests of publication bias are not valid.
Mortality

There were no studies using engagement in the intervention design that reported on child mortality.

Immunisation knowledge

We included a total of $k = 3$ studies in the analysis. The estimated average outcome based on the random-effects model was $\hat{\mu} = 0.31$ (95% CI: 0.04 to 0.58). Therefore, the average outcome differed significantly from zero ($z = 2.26, p = 0.02$), indicating a moderate significant benefit to the intervention group compared to the control group (see Supplementary Figure 58). A 95% credibility/prediction interval for the true outcomes is given by $-0.17$ to $0.79$. Hence, although the average outcome is estimated to be positive, in some studies the true outcome may in fact be negative.

According to the Q-test, the true outcomes appear to be heterogeneous ($Q(2) = 7.23, p = 0.03, \hat{\tau}^2 = 0.04, I^2 = 72.34\%$). An examination of the studentized residuals revealed that one study (Murthy et al. 2019) had a value larger than $\pm 2.39$ and may be a potential outlier in the context of this model. Indeed, sensitivity analyses leaving each study out indicated that removing Murthy and colleagues (2019) would increase the overall average effect ($\hat{\mu} = 0.45$ (95% CI: 0.25 to 0.65), and the effect would still be positive and significant ($z = 4.40, p < .001$). Removal of either of the other studies results in a non-significant effect estimate. According to the Cook’s distances, none of the studies could be considered to be overly influential. With only three studies, we were unable to test for sources of heterogeneity.

Supplementary Figure 57. Forest plot showing the observed outcomes and the estimate of the random-effects model for interventions with community engagement in the design on the intervention on childhood morbidity.

<table>
<thead>
<tr>
<th>Author (Year)</th>
<th># Participants</th>
<th>SMD [95% CI]</th>
</tr>
</thead>
<tbody>
<tr>
<td>Modi et al. (2019)</td>
<td>3470</td>
<td>0.01 [-0.06, 0.08]</td>
</tr>
<tr>
<td>Rahman et al. (2008*)</td>
<td>705</td>
<td>0.12 [-0.03, 0.27]</td>
</tr>
<tr>
<td>RE Model (Q = 1.84, df = 1, p = 0.17; $I^2 = 45.7%$)</td>
<td></td>
<td>0.05 [-0.06, 0.15]</td>
</tr>
</tbody>
</table>

* indicates some concerns or low risk of bias

<table>
<thead>
<tr>
<th>Standardized Mean Difference (SMD)</th>
</tr>
</thead>
<tbody>
<tr>
<td>-1.0</td>
</tr>
<tr>
<td>-0.5</td>
</tr>
<tr>
<td>0.0</td>
</tr>
<tr>
<td>0.5</td>
</tr>
<tr>
<td>1.0</td>
</tr>
<tr>
<td>1.5</td>
</tr>
</tbody>
</table>

BMJ Publishing Group Limited (BMJ) disclaims all liability and responsibility arising from any reliance placed on this supplemental material which has been supplied by the author(s).

Supplementray Figure 58. Forest plot showing the observed outcomes and the estimate of the random-effects model for interventions with community engagement in the design on the intervention on childhood morbidity.

**Immunisation attitudes**

Only two studies reporting on immunisation attitudes used interventions with community engagement in the intervention design. The estimated average outcome based on the random-effects model was $\hat{\mu} = -0.11$ (95% CI: −0.53 to 0.30). Therefore, the average outcome did not differ significantly from zero ($z = -0.55$, $p = 0.58$, see Supplementary Figure 59). Given the small number of studies, this result should be interpreted with caution. According to the Q-test, the true outcomes appear to be heterogeneous ($Q(1) = 4.10$, $p = 0.04$, $I^2 = 75.61\%$). With only two studies, moderator and leave-one-out analyses were not appropriate and tests of publication bias are not valid.

<table>
<thead>
<tr>
<th>Author</th>
<th># Participants</th>
<th>SMD [95% CI]</th>
</tr>
</thead>
<tbody>
<tr>
<td>Murthy et al. (2019)</td>
<td>505</td>
<td>0.11 [0.07, 0.28]</td>
</tr>
<tr>
<td>Johri et al. (2020)</td>
<td>159</td>
<td>0.34 [0.03, 0.65]</td>
</tr>
<tr>
<td>Dipeolu (2017)</td>
<td>222</td>
<td>0.53 [0.27, 0.79]</td>
</tr>
</tbody>
</table>

Supplementray Figure 58. Forest plot showing the observed outcomes and the estimate of the random-effects model for interventions with community engagement in the design on the intervention on childhood morbidity.
Vaccination card retention

No studies reporting on vaccination card retention/availability used engagement in the intervention design.

Experience and satisfaction with health services

Only one study (Engineer 2016) fell into this intervention/outcome category, thus we were unable to perform a statistical synthesis. This cluster RCT from Afghanistan did not find a significant effect of their programme on satisfaction with health services ($g = 0.20$ [95% CI: -0.11 to 0.51]), but as with the other study reporting on this outcome, it was assessed as having a high risk of bias.
Appendix 12: Community engagement in implementation autonomy of interventions

Appendix 12 presents supplementary data for analyses presented in the main body of the manuscript for the primary outcomes of community engagement in implementation autonomy, including funnel plots of publication bias and outlier analyses. It also includes a full presentation of the supplementary outcomes that were precluded from the main body of the manuscript due to word limitations.

Full immunisation

Only k = 2 studies reported on the impact of interventions with community engagement in implementation autonomy on full childhood immunisation. There was no effect on full immunisation ($\hat{\mu} = 0.23$ [95% CI: $-0.001$ to $0.47$], $z = 1.95$, $p = 0.051$). According to the $Q$-test, the true outcomes appear to be homogeneous ($Q(1) = 3.71$, $p = 0.054$, $I^2 = 73.07\%$, see Supplementary Figure 60). With only two studies, we could not test for moderation or publication bias, nor conduct sensitivity analysis by study quality or by leaving each study out.

<table>
<thead>
<tr>
<th>Author (Year)</th>
<th># Participants</th>
<th>SMD [95% CI]</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bjorkman et al. (2009* )</td>
<td>940</td>
<td>0.13 [ 0.00, 0.26]</td>
</tr>
<tr>
<td>Igarashi et al. (2010 )</td>
<td>348</td>
<td>0.37 [ 0.16, 0.59]</td>
</tr>
<tr>
<td>RE Model (Q = 3.71, df = 1, p = 0.05, $I^2 = 73.1%$)</td>
<td></td>
<td>0.23 [ 0.00, 0.47]</td>
</tr>
</tbody>
</table>

* indicates some concerns or low risk of bias

Supplementary Figure 60. Forest plot showing the observed outcomes and the estimate of the random-effects model for interventions with community engagement in implementation autonomy on full childhood immunisation.

DPT3

Only k = 3 studies examined the impact of interventions with community engagement in implementation autonomy on DPT3 vaccination. The average outcome based on the random-effects model indicated there was no effect on DPT3 vaccination ($\hat{\mu} = 0.11$ [95% CI: $-0.05$ to $0.28$], $z = 1.38$, $p = 0.17$). A 95% credibility/prediction interval for the true outcomes is given by $-0.19$ to $0.41$. Hence, although the average outcome is estimated to be positive, in some studies the true outcome may in fact be negative.

According to the $Q$-test, the true outcomes appear to be heterogeneous ($Q(2) = 10.06$, $p < 0.01$, $I^2 = 80.12\%$). An examination of the studentized residuals revealed that one study (Webster 2019) had a value larger than ±2.39 and may be a potential outlier in the context of this model. Indeed, sensitivity analyses leaving each study out indicated that removing Webster (2019) would increase the overall average effect ($\hat{\mu} = 0.20$ [95% CI: 0.10 to 0.29]), with the effect still positive and significant ($z = 3.99$, $p < .001$). The leave-one-out analysis
confirmed that neither of the other studies significantly impact the average estimate when removed. According to the Cook’s distances, none of the studies could be considered to be overly influential.

With only three studies, moderator analyses and tests of publication bias were not appropriate. With two of the three studies assessed as high risk of bias, we were also unable to conduct a sensitivity analysis by study quality.

<table>
<thead>
<tr>
<th>Author (Year)</th>
<th># Participants</th>
<th>SMD</th>
<th>[95% CI]</th>
</tr>
</thead>
<tbody>
<tr>
<td>Webster et al. (2019)</td>
<td>1404</td>
<td>-0.03</td>
<td>[-0.14, 0.07]</td>
</tr>
<tr>
<td>Björkman et al. (2009*)</td>
<td>940</td>
<td>0.20</td>
<td>[0.07, 0.32]</td>
</tr>
<tr>
<td>Shukla (2018)</td>
<td>714</td>
<td>0.20</td>
<td>[0.05, 0.34]</td>
</tr>
</tbody>
</table>

RE Model: \( Q = 10.06, df = 2, p = 0.01; \hat{I}^2 = 80.1\% \)

*indicates some concerns or low risk of bias

Supplementary Figure 61. Forest plot showing the observed outcomes and the estimate of the random-effects model for interventions with community engagement in implementation autonomy on DPT3 vaccination

Measles

Only \( k = 2 \) studies examined the impact of interventions with community engagement in implementation autonomy on measles vaccination. The average outcome based on the random-effects model indicated there was no effect on measles vaccination \( (\hat{\mu} = 0.03 \ [95\% \ CI: -0.09 \ to \ 0.15], z = 0.47, p = 0.64) \). According to the Q-test, the true outcomes appear to be homogeneous \( (Q(1) = 2.21, p = 0.14, \hat{I}^2 = 0.004, I^2 = 54.84\% \); see Supplementary Figure 62). With only two studies we were unable to test for publication bias, or complete moderator, leave-one-out, or sensitivity analyses.
Supplementary Figure 62. Forest plot showing the observed outcomes and the estimate of the random-effects model for interventions with community engagement in implementation autonomy on measles vaccination

Timeliness of full childhood immunisation

Only \( k = 2 \) studies examined the impact of interventions with community engagement in implementation autonomy on the timeliness of full childhood immunisation. The average outcome based on the random-effects model indicated there was no effect on the timeliness of full childhood immunisation (\( \hat{\mu} = 0.38 \) [95% CI: \(-0.28\) to \(1.03\)], \( z = 1.13 \), \( p = 0.26 \)). According to the \( Q \)-test, the true outcomes appear to be homogeneous (\( Q(1) = 3.57 \), \( p = 0.06 \), \( I^2 = 72.02\% \)). There was no indication of outliers in the context of this model, and with only two studies, we could not test for moderation or publication bias, or do a leave-one-out analysis. One study was assessed as having a high risk of bias (Igarashi et al. 2010) while the other was assessed as some concerns related to risk of bias (Lee 2015).

### Timeliness of Full Childhood Immunisation

<table>
<thead>
<tr>
<th>Author (Year)</th>
<th># Participants</th>
<th>SMD [95% CI]</th>
</tr>
</thead>
<tbody>
<tr>
<td>Igarashi et al. (2010)</td>
<td>348</td>
<td>0.11 [0.10, 0.32]</td>
</tr>
<tr>
<td>Lee (2015*)</td>
<td>33</td>
<td>0.80 [0.12, 1.48]</td>
</tr>
<tr>
<td>RE Model ((Q = 3.57, df = 1, p = 0.06; I^2 = 72.0%))</td>
<td></td>
<td>0.38 [0.28, 1.03]</td>
</tr>
</tbody>
</table>

*indicates some concerns or low risk of bias
Supplementary Figure 63. Forest plot showing the observed outcomes and the estimate of the random-effects model for interventions with community engagement in implementation autonomy on DPT3 vaccination timeliness

DPT3 vaccination timeliness

Only k = 2 studies examined the impact of interventions with community engagement in implementation autonomy on DPT3 vaccination timeliness. The average outcome based on the random-effects model indicated there was no effect on the timeliness of DPT3 vaccination ($\hat{\mu} = 0.04$ [95% CI: $-0.06$ to $0.13$], $z = 0.75$, $p = 0.45$). According to the $Q$-test, the true outcomes appear to be homogeneous ($Q(1) = 0.63$, $p = 0.45$, $\hat{\tau}^2 = 0.00$, $I^2 = 0.00\%$). There was no indication of outliers in the context of this model, and with only two studies, we could not test for moderation or publication bias, or conduct a leave-one-out analysis. Both studies were assessed as having a high risk of bias.

Supplementary Figure 64. Forest plot showing the observed outcomes and the estimate of the random-effects model for interventions with community engagement in implementation autonomy on DPT3 vaccination timeliness

Measles vaccination timeliness

No studies reporting on the timeliness of measles vaccinations used community engagement in implementation autonomy.

Secondary outcomes

Partial immunisation
Only one study (Webster et al. 2019) fell into this intervention/outcome category, thus we were unable to perform a statistical synthesis. This cluster RCT from Uganda found a null effect of their programme on partial childhood immunisation ($g = 0.02 \ [95\% \ CI: -0.08 \ to \ 0.13]$), but like most studies, it was assessed as having a high risk of bias.

**BCG**

We included $k = 2$ studies in this analysis. The estimated average outcome was $\hat{\mu} = 0.03 \ [95\% \ CI: -0.05 \ to \ 0.11]$, $z = 0.75$, $p = 0.46$ indicating no difference between the treated group and the untreated group (Supplementary Figure 62). According to the $Q$-test, the true outcomes appear to be homogeneous ($Q(1) = 0.37$, $p = 0.54$, $I^2 = 0.00\%$). With only two studies and no heterogeneity, moderator and leave-one-out analyses were not appropriate and tests of publication bias are not valid.

![Forest plot showing the observed outcomes and the estimate of the random-effects model for interventions with community engagement in implementation autonomy on BCG vaccination](image)

Supplementary Figure 65. Forest plot showing the observed outcomes and the estimate of the random-effects model for interventions with community engagement in implementation autonomy on BCG vaccination

**DPT1**

Only one study (Webster et al. 2019) fell into this intervention/outcome category, thus we were unable to perform a statistical synthesis. This cluster RCT from Uganda found a null effect of their programme on DPT1 vaccination ($g = 0.04 \ [95\% \ CI: -0.07 \ to \ 0.14]$), but like most studies, it was assessed as having a high risk of bias.

**DPT2**

Only one study (Webster et al. 2019) fell into this intervention/outcome category, thus we were unable to perform a statistical synthesis. This cluster RCT from Uganda found a null effect of their programme on DPT1 vaccination ($g = 0.01 \ [95\% \ CI: -0.10 \ to \ 0.11]$), but like most studies, it was assessed as having a high risk of bias.
**OPV0**

Only one study (Webster et al. 2019) fell into this intervention/outcome category, thus we were unable to perform a statistical synthesis. This cluster RCT from Uganda found a null effect of their programme on OPV0 vaccination ($g = 0.01$ [95% CI: -0.10 to 0.11]), but like most studies, it was assessed as having a high risk of bias.

**OPV1**

Only one study (Webster et al. 2019) fell into this intervention/outcome category, thus we were unable to perform a statistical synthesis. This cluster RCT from Uganda found a null effect of their programme on OPV1 vaccination ($g = 0.03$ [95% CI: -0.07 to 0.14]), but like most studies, it was assessed as having a high risk of bias.

**OPV2**

Only one study (Webster et al. 2019) fell into this intervention/outcome category, thus we were unable to perform a statistical synthesis. This cluster RCT from Uganda found a null effect of their programme on OPV2 vaccination ($g = 0.04$ [95% CI: -0.06 to 0.15]), but again, it was assessed as having a high risk of bias.

**OPV3**

Only two studies examined the impact of interventions with community engagement in implementation autonomy on OPV3 vaccination. The estimated average outcome was $\hat{\mu} = 0.03$ [95% CI: $-0.12$ to $0.18$], which was not significant ($z = 0.35$, $p = 0.73$), indicating there was no difference between the intervention group and the control group. According to the $Q$-test, the true outcomes appear to be homogeneous ($Q(1) = 3.38$, $p = 0.07$, $\hat{\tau}^2 = 0.01$, $I^2 = 70.38\%$; see Supplementray Figure 66). With only two studies we were unable to test for publication bias or sources of heterogeneity or complete a leave-one-out analysis.

<table>
<thead>
<tr>
<th>Author (Year)</th>
<th># Participants</th>
<th>SMD [95% CI]</th>
</tr>
</thead>
<tbody>
<tr>
<td>Webster et al. (2019)</td>
<td>1404</td>
<td>-0.05 [-0.15, 0.06]</td>
</tr>
<tr>
<td>Bjorkman et al. (2009*)</td>
<td>940</td>
<td>0.11 [-0.02, 0.24]</td>
</tr>
</tbody>
</table>

*indicates some concerns or low risk of bias

Supplementary Figure 66. Forest plot showing the observed outcomes and the estimate of the random-effects model for interventions with community engagement in implementation autonomy on OPV3 vaccinations.
Dropouts

Only one study (Webster et al. 2019) fell into this intervention/outcome category, thus we were unable to perform a statistical synthesis. This cluster RCT from Uganda found a small but significant negative impact of their programme on full childhood immunisation ($g = -0.24$ [95% CI: $-0.35$ to $-0.14$]), but like most studies, it was assessed as having a high risk of bias.

Morbidity

We used reports of diarrhea (most typically in the past two weeks) as a proxy for childhood morbidity. In all cases, effects were reverse coded such that positive effects always indicate a benefit to the treated group. Thus a positive effect here would be interpreted as a reduction in diarrhea among treated participants compared to control participants. Only one study (Webster et al. 2019) fell into this intervention/outcome category, thus we were unable to perform a statistical synthesis. This cluster RCT from Uganda found a null effect of their programme on morbidity ($g = 0.06$ [95% CI: $-0.02$ to $0.13$]), but like most studies, it was assessed as having a high risk of bias.

Mortality

There were no studies using community engagement in implementation autonomy that reported on child mortality.

Immunisation knowledge

There were no studies reporting on immunisation knowledge that used community engagement in implementation autonomy.

Immunisation attitudes

There were no studies reporting on immunisation attitudes that used interventions with community engagement in implementation autonomy.

Vaccination card retention

Only one study (Webster et al. 2019) fell into this intervention/outcome category, thus we were unable to perform a statistical synthesis. This cluster RCT from Uganda found a very small positive but non-significant effect of their programme on full childhood immunisation ($g = 0.04$ [95% CI: $-0.06$ to $-0.15$]), but like most studies, it was assessed as having a high risk of bias.

Experience and satisfaction with health services

No studies reporting on experience and satisfaction with health services used engagement in implementation autonomy.

Formal health worker’s motivation, capacity and performance

164
No studies reporting on formal health worker’s motivation, capacity and performance used engagement in implementation autonomy.
Appendix 13: Multiple engagement types

**Full immunisation**

We included k = 9 studies examining the impact of interventions with multiple engagement types on full childhood immunisation. The average outcome based on the random-effects model indicated that there was no effect on full childhood immunisation ($\hat{\mu} = 0.22$ [95% CI: $-0.12$ to $0.56$], $z = 1.27$, $p = 0.20$). A 95% credibility/prediction interval for the true outcomes is given by $-0.81$ to $1.25$. Hence, although the average outcome is estimated to be positive, in some studies the true outcome may in fact be negative.

According to the $Q$-test, the true outcomes appear to be heterogeneous ($Q(8) = 388.56$, $p < 0.01$, $I^2 = 97.94\%$, see Supplementary Figure 67). Leave-one-out analysis indicated that the removal of Banerjee et al. (2010) would reduce the average estimate to .03, but it would still be positive and non-significant. There were no high or moderate quality studies, so we were unable to perform a sensitivity analysis. None of the moderators were significant sources of heterogeneity in the context of this model, including whether the specific combination of engagement types led to different effects (see Supplementary Table 3), but for full childhood immunisation we found no differences by engagement packages ($\hat{\beta} = 0.36$, $p = 0.30$ [95% CI: $-0.31$ to $1.02$]).

**DPT3**

We included k = 7 studies examining the impact of interventions with multiple engagement types on DPT3 vaccination. The average outcome based on the random-effects model indicated that there was a significant effect on DPT3 vaccination ($\hat{\mu} = 0.20$ [95% CI: $0.06$ to $0.34$], $z = 2.85$, $p < 0.01$). A 95% credibility/prediction interval for the true outcomes is given by $-0.14$ to $0.55$. Hence, although the average outcome is estimated to be positive, in some studies the true outcome may in fact be negative.
According to the $Q$-test, the true outcomes appear to be heterogeneous ($Q(6) = 49.71, p < 0.01, I^2 = 87.93\%$). An examination of the studentized residuals revealed that none of the studies had a value larger than \pm 2.69 and hence there was no indication of outliers in the context of this model. This was confirmed with a leave-one-out analysis. According to the Cook’s distances, none of the studies could be considered to be overly influential. With no high or moderate quality studies, we were unable to perform a sensitivity analysis. There were several moderators that were significant sources of heterogeneity. Evaluation period was significant such that each additional month between the end of the intervention and the collection of outcome data reduced the size of the effect by .01 standard deviation units ($\hat{\beta} = -0.01, p = 0.003 [95\% CI: -0.02 to -0.004]$), suggesting smaller long term effect of the interventions. Publication year was also significant, such that each additional year reduced the size of the effect by .03 standard deviation units ($\hat{\beta} = -0.03, p = 0.03 [95\% CI: -0.05 to -0.003]$). In other words, more recent studies have found smaller effects. Finally, baseline DPT3 coverage rates were significant such that a one unit increase in baseline DPT3 coverage was associated with a decrease of .56 in the effect of the programme ($\hat{\beta} = -0.56, p = 0.002 [95\% CI: -0.90 to -0.21]$). In other words, the programmes were significantly more effective in areas with lower baseline coverage rates. We also tested whether the specific combination of engagement types led to different effects, but for DPT3 vaccination we found no differences by engagement packages ($\hat{\beta} = -0.07, p = 0.66 [95\% CI: -0.40 to -0.25]$).

**Supplementary Figure 68. Forest plot showing the observed outcomes and the estimate of the random-effects model for interventions with multiple engagement types on DPT3 vaccination**

**Measles**

We included $k = 6$ studies examining the impact of interventions with multiple engagement types on measles vaccination. The average outcome based on the random-effects model indicated that there was no effect on measles coverage ($\hat{\mu} = 0.03 [95\% CI: -0.10 to 0.16], z = 0.44, p = 0.66$). A 95% credibility/prediction interval for the true outcomes is given by $-0.26$ to $0.31$. Hence, although the average outcome is estimated to be positive, in some studies the true outcome may in fact be negative.
According to the \( Q \)-test, the true outcomes appear to be heterogeneous (\( Q(5) = 37.76, p < 0.01, I^2 = 86.76\% \); see Supplementary Figure 69). An examination of the studentized residuals revealed that one study (Sankar 2013) had a value larger than \( \pm 2.64 \) and may be a potential outlier in the context of this model. According to the Cook’s distances, none of the studies could be considered to be overly influential. Indeed, sensitivity analyses leaving each study out indicated that removing Sankar (2013) would increase the overall average effect (\( \hat{\mu} = 0.09 \) (95% CI: -0.02 to 0.19), but the effect would still be non-significant (\( z = 1.58, p = 0.11 \)), and that no other study’s removal would significantly impact the average estimated effect. With only one high or medium quality study, we were unable to conduct sensitivity analysis by study quality. None of the moderators that could be tested were significant in the context of this model (see Supplementary Table 3). For measles vaccinations, we were unable to test whether different combinations of community engagement produced different results because all but one study used a combination of engagement in the implementation and engagement as the intervention.

**Supplementary Figure 69. Forest plot showing the observed outcomes and the estimate of the random-effects model for interventions with multiple engagement types on measles vaccination**

**Timeliness of full childhood immunisation**

There were no studies using multiple engagement types that reported on the timeliness of full childhood immunisation.

**Timeliness of DPT3 vaccination**

Only one study reporting on timeliness of DPT3 vaccination used an intervention with multiple engagement types (Oyo-Ita et al. 2020). This RCT conducted in Nigeria found a small but not-significant positive effect (\( g = 0.10 \) [95%CI; -0.2 to 0.23]. This study was assessed as having a high risk of bias.

**Timeliness of measles vaccination**

Finally, only one study reporting on timeliness of measles vaccination used an intervention with multiple engagement types (Oyo-Ita et al. 2020). This RCT conducted in Nigeria found a moderate and significant positive
effect (g = 0.25 [95% CI; 0.13 to 0.36]). They used a combination of engagement in implementation autonomy and engagement as the intervention. This study was also assessed as having a high risk of bias.

**Secondary outcomes**

**Partial immunisation**

We included a total of \( k = 4 \) studies in the analysis. The estimated average outcome based on the random-effects model was \( \hat{\mu} = 0.28 \) (95% CI: −0.01 to 0.56). Therefore, the average outcome did not differ significantly from zero \( (z = 1.91, p = 0.06) \), indicating no difference between the intervention and control groups (see Supplementary Figure 70). A 95% credibility/prediction interval for the true outcomes is given by −0.35 to 0.91. Hence, although the average outcome is estimated to be positive, in some studies the true outcome may in fact be negative.

According to the \( Q \)-test, the true outcomes appear to be heterogeneous \( (Q(3) = 101.75, p < 0.01, \hat{t}^2 = 0.08, I^2 = 97.05\%) \). An examination of the studentized residuals revealed that none of the studies had a value larger than ±2.50 and hence there was no indication of outliers in the context of this model. However, a leave-one-out analysis did indicate that removing Herrera-Almanza & Rosales-Rueda (2018) would result in a larger and statistically significant effect \( (\hat{\mu} = 0.37, [95\% CI: −0.01 to 0.56], p = .03) \). According to the Cook’s distances, none of the studies could be considered to be overly influential. Of the moderators we were able to test, none were significant sources of heterogeneity (see Supplementary Table 3). We were unable to test for differences among engagement packages because only one study used a combination of engagement in the design and engagement as the intervention (Gurley et al. 2020) while the remaining studies used a combination of engagement in implementation autonomy and engagement as the intervention.

<table>
<thead>
<tr>
<th>Author (Year)</th>
<th># Participants</th>
<th>SMD [95% CI]</th>
</tr>
</thead>
<tbody>
<tr>
<td>Herrera-Almanza &amp; Rosales-Rueda (2018*)</td>
<td>4612</td>
<td>0.02 [-0.04, 0.08]</td>
</tr>
<tr>
<td>Gurley et al. (2020)</td>
<td>1243</td>
<td>0.04 [-0.07, 0.16]</td>
</tr>
<tr>
<td>Sankar (2013)</td>
<td>1056</td>
<td>0.46 [0.34, 0.58]</td>
</tr>
<tr>
<td>Oye-Iita et al. (2020)</td>
<td>1716</td>
<td>0.60 [0.48, 0.72]</td>
</tr>
</tbody>
</table>

**RE Model** \( (Q = 101.75, df = 3, p = 0.00, \hat{t}^2 = 97.1\%) \)

\*indicates some concerns or low risk of bias

Supplementary Figure 70. Forest plot showing the observed outcomes and the estimate of the random-effects model for interventions with multiple engagement types on partial immunisation.

**BCG**
We included a total of \( k = 4 \) studies in the analysis. The estimated average outcome based on the random-effects model was \( \hat{\mu} = 0.22 \) (95\% CI: \(-0.07\) to \(0.52\)). Therefore, the average outcome did not differ significantly from zero \((z = 1.48, p = 0.14)\), indicating no differences between the treatment and control groups (see Supplementary Figure 71). A 95\% credibility/prediction interval for the true outcomes is given by \(-0.41\) to \(0.86\). Hence, although the average outcome is estimated to be positive, in some studies the true outcome may in fact be negative.

According to the \( Q \)-test, the true outcomes appear to be heterogeneous \((Q(3) = 75.36, p < 0.01, I^2 = 96.02\% \)) . An examination of the studentized residuals revealed that one study (Banerjee 2010) had a value larger than \( \pm 2.50 \) and may be a potential outlier in the context of this model. Indeed, sensitivity analyses leaving each study out indicated that removing Banerjee (2010) would reduce the overall average effect \( (\hat{\mu} = 0.05 \) (95\% CI: \(-0.13\) to \(0.23\)), but the effect is still positive and non-significant \((z = 0.56, p = 0.58)\). The leave-one-out analysis did not indicate that leaving any other studies out would significantly change the average estimated effect. According to the Cook’s distances, none of the studies could be considered to be overly influential. With all studies assessed as high risk of bias, we were unable to conduct sensitivity analysis by study quality. We tested all moderators and only post-intervention versus change from baseline was a significant predictor of BCG vaccination, such that studies examining change from baseline had higher effects than studies examining post-intervention changes by \( .55 \) standard deviation units \((\hat{\beta} = -0.55 [95\% CI: -0.69 to -0.404], p < .001)\). We could not test for moderation by engagement components because all but one study used a combination of engagement in implementation autonomy and engagement as the intervention.

**Supplementary Figure 71.** Forest plot showing the observed outcomes and the estimate of the random-effects model for interventions with multiple engagement types on BCG vaccination.

**DPT1**

Only two studies reporting on DPT1 vaccination used interventions with multiple engagement types. The estimated average outcome based on the random-effects model was \( \hat{\mu} = -0.17 \) (95\% CI: \(-0.29\) to \(-0.05\)), \( z = -2.81, p = 0.005 \), thus there was a significant negative impact of the programmes on DPT1 vaccination, (see Supplementary
Figure 72). Given the small number of studies, this result should be interpreted with caution. According to the Q-test, the true outcomes appear to be homogeneous ($Q(1) = 0.03, p = 0.87$, $\hat{t}^2 = 0.00$, $I^2 = 0.00\%$). With only two studies, moderator and leave-one-out analyses were not appropriate and tests of publication bias are not valid.

![Forest plot showing the observed outcomes and the estimate of the random-effects model for interventions with multiple engagement types on DPT1 vaccination.](image)

*Supplementary Figure 72. Forest plot showing the observed outcomes and the estimate of the random-effects model for interventions with multiple engagement types on DPT1 vaccination.*

**DPT2**

Only one study (Sankar 2013) fell into this intervention/outcome category, thus we were unable to perform a statistical synthesis. This quasi-experimental study from India found a null effect of their programme on OPV0 vaccination ($g = 0.10$ [95% CI: -0.02 to 0.22]), but like most studies, it was assessed as having a high risk of bias. Their programme included a combination of engagement in the implementation and engagement as the intervention.

**OPV0**

Only one study (Sankar 2013) fell into this intervention/outcome category, thus we were unable to perform a statistical synthesis. This quasi-experimental study from India found a significant positive effect of their programme on OPV0 vaccination ($g = 0.43$ [95% CI: 0.31 to 0.55]), but like most studies, it was assessed as having a high risk of bias. Their programme included a combination of engagement in implementation autonomy and engagement as the intervention.

**OPV1**

Only two studies reporting on OPV1 vaccination used interventions with multiple engagement types. The estimated average outcome based on the random-effects model was $\hat{\mu} = 0.22$ (95% CI: −0.15 to 0.59). Therefore, the average
outcome did not differ significantly from zero \((z = 1.17, p = 0.24, \text{see Supplementary Figure 73})\). Given the small number of studies, this result should be interpreted with caution. According to the Q-test, the true outcomes appear to be homogeneous \((Q(1) = 3.61, p = 0.06, \hat{\tau}^2 = 0.05, I^2 = 72.32\%)\). With only two studies, moderator and leave-one-out analyses were not appropriate and tests of publication bias are not valid.

<table>
<thead>
<tr>
<th>Author (Year)</th>
<th># Participants</th>
<th>SMD [95% CI]</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sankar (2013)</td>
<td>1056</td>
<td>0.07 [0.05, 0.19]</td>
</tr>
<tr>
<td>Admassie et al. (2009)</td>
<td>106</td>
<td>0.46 [0.08, 0.84]</td>
</tr>
</tbody>
</table>

RE Model \((\hat{\mu} = 3.61, df = 1, p = 0.06, \hat{I}^2 = 72.3\%)\)

* indicates some concerns or low risk of bias

Supplementary Figure 73. Forest plot showing the observed outcomes and the estimate of the random-effects model for interventions with multiple engagement types on OPV1 vaccination.

**OPV2**

Only two studies reporting on OPV2 vaccination used interventions with multiple engagement types. The estimated average outcome based on the random-effects model was \(\hat{\mu} = 0.34\) \((95\% \text{ CI: 0.19 to 0.50, } z = 4.28, p < 0.001)\), indicating a small but significant benefit to the treated group compared to the control participants (see Supplementary Figure 74). Given the small number of studies, this result should be interpreted with caution. According to the Q-test, the true outcomes appear to be homogeneous \((Q(1) = 1.18, p = 0.28, \hat{\tau}^2 = 0.004, I^2 = 15.51\%)\). With only two studies, moderator and leave-one-out analyses were not appropriate and tests of publication bias are not valid. Both studies used a combination of engagement in implementation autonomy and engagement as the intervention.
Supplementary Figure 74. Forest plot showing the observed outcomes and the estimate of the random-effects model for interventions with multiple engagement types on OPV2 vaccination.

OPV3

We included a total of $k = 3$ studies in the analysis. The estimated average outcome based on the random-effects model was $\hat{\mu} = 0.48$ ([95% CI: −0.24 to 1.20], $z = 1.30$, $p = 0.19$), indicating no difference between the intervention group and the control group (see Supplementary Figure 75). A 95% credibility/prediction interval for the true outcomes is given by −0.95 to 1.91. Hence, although the average outcome is estimated to be positive, in some studies the true outcome may in fact be negative.

According to the $Q$-test, the true outcomes appear to be heterogeneous ($Q(2) = 185.96$, $p < 0.001$, $I^2 = 98.92\%$). An examination of the studentized residuals revealed that none of the studies had a value larger than ±2.39 and hence there was no indication of outliers in the context of this model. However, the leave-one-out analysis indicates that removing Okeke et al. (2017) would result in a larger and statistically significant effect estimate ($\hat{\mu} = 0.76$ [95% CI: 0.37 to 1.16], $z = 3.80$, $p < 0.001$). According to the Cook’s distances, none of the studies could be considered to be overly influential. With only three studies contributing effects, we were unable to examine sources of heterogeneity.
Supplementary Figure 75. Forest plot showing the observed outcomes and the estimate of the random-effects model for interventions with multiple engagement types on OPV3 vaccination.

**Dropouts**

Only two studies reporting on immunisation dropouts used interventions with multiple engagement types. All effects were computed such that a positive effect reflects a benefit to the treated group (e.g. a reduction in dropouts). The estimated average outcome based on the random-effects model was $\hat{\mu} = 0.18$ (95% CI: 0.06 to 0.30). Therefore, the average outcome differed significantly from zero, indicating a small but significant benefit to the treated group compared to the control group ($z = 2.92$, $p = 0.004$, see Supplementary Figure 76). Given the small number of studies, this result should be interpreted with caution. According to the Q-test, the true outcomes appear to be homogeneous ($Q(1) = 2.35$, $p = 0.13$, $\hat{\tau}^2 = 0.004$, $I^2 = 57.38\%$). With only two studies, moderator and leave-one-out analyses were not appropriate and tests of publication bias are not valid.
Supplementary Figure 76. Forest plot showing the observed outcomes and the estimate of the random-effects model for interventions with multiple engagement types on vaccination dropouts.

**Morbidity**

We used reports of diarrhea (most typically in the past two weeks) as a proxy for childhood morbidity. In all cases, effects were reverse coded such that positive effects always indicate a benefit to the treated group. Thus a positive effect here would be interpreted as a reduction in diarrhea among treated participants compared to control participants. Only two studies reporting on childhood morbidity used interventions with multiple engagement types. The estimated average outcome based on the random-effects model was $\hat{\mu} = -0.10$ (95% CI: $-0.60$ to $0.40$). Therefore, the average outcome did not differ significantly from zero, indicating no difference between the treated group compared to the control group ($z = -0.38$, $p = 0.70$, see Supplementary Figure 77). Given the small number of studies, this result should be interpreted with caution. According to the Q-test, the true outcomes appear to be heterogeneous ($Q(1) = 5.23$, $p = 0.02$, $I^2 = 57.4\%$). With only two studies, moderator and leave-one-out analyses were not appropriate and tests of publication bias are not valid.
Mortality

We included a total of $k = 3$ studies in the analysis. The observed outcomes ranged from $-0.25$ to $-0.002$. The estimated average outcome based on the random-effects model was $\hat{\mu} = -0.04$ (95% CI: $-0.10$ to $0.03$, see Supplementary Figure 78). Therefore, the average outcome did not differ significantly from zero ($z = -1.15$, $p = 0.25$), indicating no difference between the intervention participants and control participants on child mortality. A 95% credibility/prediction interval for the true outcomes is given by $-0.14$ to $0.07$. Hence, although the average outcome is estimated to be negative, in some studies the true outcome may in fact be positive.

The $Q$-test for heterogeneity was not significant, but some heterogeneity may still be present in the true outcomes ($Q(2) = 5.37$, $p = 0.02$, $I^2 = 80.9\%$). An examination of the studentized residuals revealed that none of the studies had a value larger than ±2.39 and hence there was no indication of outliers in the context of this model. This was confirmed by a leave-one-out analysis. According to the Cook’s distances, none of the studies could be considered to be overly influential. With only three studies, moderator analyses were not appropriate and tests of publication bias are not valid.
Only two studies reporting on immunisation knowledge used interventions with multiple engagement types. The estimated average outcome based on the random-effects model was $\hat{\mu} = 0.09$ (95% CI: 0.04 to 0.13). Therefore, the average outcome differed significantly from zero, indicating a small but significant benefit to the treated group compared to the control group ($z = 3.55$, $p < 0.001$), see Supplementary Figure 79). Given the small number of studies, this result should be interpreted with caution. According to the Q-test, the true outcomes appear to be homogeneous ($Q(1) = 0.19$, $p = 0.66$, $\tau^2 = 0.00$, $I^2 = 0.00\%$). With only two studies, moderator and leave-one-out analyses were not appropriate and tests of publication bias are not valid.

---

Supplementary Figure 78. Forest plot showing the observed outcomes and the estimate of the random-effects model for interventions with multiple engagement types on childhood morbidity.

**Immunisation knowledge**

Only two studies reporting on immunisation knowledge used interventions with multiple engagement types. The estimated average outcome based on the random-effects model was $\hat{\mu} = 0.09$ (95% CI: 0.04 to 0.13). Therefore, the average outcome differed significantly from zero, indicating a small but significant benefit to the treated group compared to the control group ($z = 3.55$, $p < 0.001$), see Supplementary Figure 79). Given the small number of studies, this result should be interpreted with caution. According to the Q-test, the true outcomes appear to be homogeneous ($Q(1) = 0.19$, $p = 0.66$, $\tau^2 = 0.00$, $I^2 = 0.00\%$). With only two studies, moderator and leave-one-out analyses were not appropriate and tests of publication bias are not valid.
Supplementary Figure 79. Forest plot showing the observed outcomes and the estimate of the random-effects model for interventions with multiple engagement types on immunisation knowledge.

Immunisation attitudes

Only two studies reporting on immunisation attitudes used interventions with multiple engagement types. The estimated average outcome based on the random-effects model was $\mu = 0.06$ (95% CI: 0.01 to 0.11). Therefore, the average outcome differed significantly from zero, indicating a small but significant benefit to the treated group compared to the control group ($z = 2.50$, $p = 0.01$), see Supplementary Figure 80). Given the small number of studies, this result should be interpreted with caution. According to the Q-test, the true outcomes appear to be homogeneous ($Q(1) = 0.50$, $p = 0.48$, $I^2 = 0.00$, $I^2 = 0.00\%$). With only two studies, moderator and leave-one-out analyses were not appropriate and tests of publication bias are not valid.
Only two studies reporting on vaccination card availability/retention used interventions with multiple engagement types. The estimated average outcome based on the random-effects model was $\hat{\mu} = -0.01$ (95% CI: $-0.05$ to $0.02$). Therefore, the average outcome did not differ significantly from zero ($z = -0.78$, $p = 0.43$, see Supplementary Figure 81). Given the small number of studies, this result should be interpreted with caution. According to the Q-test, the true outcomes appear to be homogeneous ($Q(5) = 0.13$, $p = 0.43$, $I^2 = 0.00$, $I^2 = 0.00$); With only two studies and no heterogeneity, moderator and leave-one-out analyses were not appropriate and tests of publication bias are not valid.
Supplementary Figure 81. Forest plot showing the observed outcomes and the estimate of the random-effects model for interventions with multiple engagement types on vaccination card availability
Appendix 14: Qualitative synthesis of results

Barriers to immunisation

Most information on barriers came from studies which used engagement as the intervention. Often, multiple barriers were cited in the same sentence. The vast majority of the barriers were identified through individual interviews and focus group discussions in the primary papers from which they were drawn. In addition, these were often supported by the authors’ reflections, implementers’ experiences, and observations. Sensitivity analysis, in which only qualitative papers that received an assessment score greater than 20 were included, indicated similar results. Therefore, we view the strength of the evidence to be high. Behavioural, social and practical barriers faced by caregivers were more consistently cited than direct constraints within the health system (Supplementary Figure 82, Supplementary Table 4).

Supplementary Figure 82: Summary figure: Behavioural, social and practical barriers faced by caregivers were more consistently experienced than constraints in delivery of immunisation services

The size of the rectangles reflects the number of times a unique theme was mentioned by an IE and its associated papers. Larger rectangles indicate that a theme was mentioned more times. Nested boxes reflect the proportion of times the main theme text also fell into the sub-them.

Supplementary Table 4: Comparison of the barriers to immunisation in communities in which community engagement activities were undertaken.

<table>
<thead>
<tr>
<th>Community engagement as the intervention</th>
<th>Community engagement in the design of the intervention</th>
<th>Community engagement in the implementation autonomy of the intervention</th>
</tr>
</thead>
</table>

1 The hierarchy charts showing barriers by type of community engagement are in Appendix H: barriers to immunisation
<table>
<thead>
<tr>
<th>Availability of resources</th>
<th>Staff</th>
<th>Stockouts</th>
<th>Post abandonment</th>
</tr>
</thead>
<tbody>
<tr>
<td>Absenteeism</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Stockouts</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Building infrastructure</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Quality of services</td>
<td>Attitudes of health workers</td>
<td>Attitudes of health workers</td>
<td>Attitudes of health workers</td>
</tr>
<tr>
<td></td>
<td>Quality affected attendance</td>
<td>Services not implemented as intended</td>
<td></td>
</tr>
<tr>
<td>Costs</td>
<td>Other costs</td>
<td>Other costs</td>
<td>Other costs</td>
</tr>
<tr>
<td></td>
<td>Cost of the vaccine*</td>
<td></td>
<td>Cost of the vaccine</td>
</tr>
<tr>
<td>Logistics or distance</td>
<td>Distance</td>
<td>Convenience</td>
<td>Convenience</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Distance and access</td>
<td>Transport</td>
</tr>
<tr>
<td>Social norms</td>
<td>Religious beliefs</td>
<td>Spousal consent</td>
<td>Religious beliefs</td>
</tr>
<tr>
<td></td>
<td>Beliefs of others</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Fear</td>
<td>Side effects (and</td>
<td>Side effects (from past experience)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>overcoming this fear)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Trust in the government</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>or medicine</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Understanding of importance</td>
<td>Cost and importance</td>
<td>Fear and importance</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Behavioural, social and practical barriers faced by caregivers**

The most consistent behavioural, social or practical barrier faced by caregivers was cost. However, the kind of cost which posed a barrier was inconsistent, and mostly not that of the vaccination itself. Associated costs include opportunity costs, transport costs, and the costs of treating side effects. The cost of the immunisation itself was mentioned as a barrier by community members that received engagement as the intervention or engagement in the implementation of the intervention.

*Logistics and distance* were also consistent barriers to immunisation in settings where all three types of interventions occurred. The nature of the logistical or distance challenge was variable. Distance alone was reported as an issue in the communities with engagement as the intervention. However, Olken 2014 (World bank 2018) mention other logistical barriers such as wait times and language barriers. In settings where the community was involved in the design or implementation of the intervention, convenience was an issue in addition to distance. Community members in these areas reported issues related to wait times and the timing of clinic visits. Adamu 2019 (Adamu 2019 pre-implementation and Adamu 2019 Dissertation) mention demand being lower than it could be because women were unable to arrive at clinics at specified times due to competing demands.

*Fear* of side effects was common in communities in which engagement was the intervention or engagement was used in the design of the intervention. The concern was consistently related to known side effects and affected by past experience of side effect and trust in the government or medicine. The issue of vaccines causing sterilization was also mentioned. In one case, the intervention reduced these fears (Gurley 2020 (Gurley 2020)).

*Social norms* around immunisation posed consistent barriers; however, the source of the social restriction was variable. Spousal consent limited immunisations in areas where engagement was used in the implementation of interventions;
women could not vaccinate their children without the explicit approval of their husband, who may refuse. Religious beliefs sometimes resulted in the preference for non-Western medicine. In areas that received engagement as the intervention, religious beliefs and the influence of beliefs held by others sometimes inhibited immunisation. In particular, mother-in-laws and other family members who do not value immunisation may pose a barrier to immunisations.

Failure to understand the importance of vaccination was not consistently observed but reported in settings that received engagement as the intervention. When this barrier was mentioned, it was always paired with another, such as cost or fear. Authors implied that these other barriers were experienced and could not be overcome because there was a lack of understanding of the importance of vaccinations.

**Constraints in delivery of immunisation services**

Constraints in the delivery of immunisation services were widely reported but did not seem to be as common of a barrier as the behavioural, social and practical barriers faced by caregivers. The availability of services was the most consistently reported delivery barrier across all three types of engagement. Limited staff and stockouts were dominant delivery challenges. Communities that had engagement as the intervention or were engaged in implementation consistently have issues with staff being assigned but unavailable. Poor infrastructure, such as dilapidated buildings, was a challenge in communities that were involved in the design of the intervention. In areas where engagement was the intervention or the community was involved in the design of the intervention, the quality of services was also a more consistently reported barrier than any behavioural, social, or practical barrier faced by caregivers. Often, these were related to health worker attitudes, either a lack of motivation or conflict with caregivers.

**Barriers to immunisation by type of community engagement**

![Figure 83: Engagement as intervention](image-url)
Supplementary Table 5: Characteristic quotes of themes identified related to barriers to immunisation for interventions were community engagement.
<table>
<thead>
<tr>
<th>Primary paper</th>
<th>Paper from which citation comes</th>
<th>Theme demonstrated</th>
<th>Quote</th>
<th>Page number</th>
</tr>
</thead>
<tbody>
<tr>
<td>Financial</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Okeke et al 2017</td>
<td>Okeke et al 2017</td>
<td>Cost of vaccination</td>
<td>costs associated with use of services and lack of adequate transportation, hindered uptake.</td>
<td>54</td>
</tr>
<tr>
<td>Andersson 2009</td>
<td>Mitchell 2009</td>
<td>Other costs</td>
<td>Yet the groups noted that immediate financial and time costs associated with having a child vaccinated weigh very heavily in comparison with the potential costs associated with measles in the future.</td>
<td>7</td>
</tr>
<tr>
<td>Social norms</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Banerjee et al 2020</td>
<td>Appendix K</td>
<td>Religion</td>
<td>Religion emerged as an implicit theme. Only two respondents mentioned explicitly that religious belief might inhibit immunisation: that God would decide the fate of their child.</td>
<td>7</td>
</tr>
<tr>
<td>Andersson 2009</td>
<td>Cockroft 2009</td>
<td>Beliefs of others</td>
<td>“My mother-in-law says ‘What kind of children have you produced that they ought to be vaccinated? We were never vaccinated, so why are you behaving so delicately?”” (female group, Khanewal)</td>
<td>8</td>
</tr>
<tr>
<td>Understanding of importance</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Olken 2014</td>
<td>Grayman 2013</td>
<td>Costs and importance</td>
<td>Typically the answer is that the parents are too busy working on their farms, or they feel that posyandu is no longer necessary after they are older than two years of age.</td>
<td>66</td>
</tr>
<tr>
<td>Paramanik 2018</td>
<td>Paramanik 2020</td>
<td>Fear and importance</td>
<td>As reported by ASHAs the mothers-in-law were not able to accept the concept of vaccination. They apparently argued on the ground that they had children too and that they never took them for any vaccination, that post vaccination the child gets fever which causes sleep disturbance for the child and in turn negatively affects the whole family.</td>
<td>84</td>
</tr>
</tbody>
</table>
### Logistics or distance

<table>
<thead>
<tr>
<th>Author</th>
<th>Year</th>
<th>Distance</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Alhassan</td>
<td>2019</td>
<td></td>
<td>In Ghana, access to maternal and child health services is impeded by longer travel times to health facilities.</td>
</tr>
<tr>
<td>Oyo-Ita</td>
<td>2020</td>
<td></td>
<td>children whose mothers found it difficult to reach the health facility</td>
</tr>
</tbody>
</table>

### Fear

<table>
<thead>
<tr>
<th>Author</th>
<th>Year</th>
<th>Topic</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Banerjee</td>
<td>2020</td>
<td>Appendix K Trust</td>
<td>A couple of respondents suggested this could be interacting with a lack of trust of front line health workers.</td>
</tr>
<tr>
<td>Oyo-Ita</td>
<td>2020</td>
<td>Oyo-Ita 2020 Side effects</td>
<td>She said she doesn’t want anybody to give her child injection so that the child will not become sick</td>
</tr>
<tr>
<td>Gurley</td>
<td>2020</td>
<td>Gurley 2020 Overcoming fear of side effects</td>
<td>We observed changes in how FGD respondents discussed side effects over the course of the study, from a perception that they were severe and not worth the risk of vaccination, to a sense they could be handled and are not a reason to forgo vaccination. I also took my child for vaccination. She cried for two days and I said let her cry; at least she will be safe in future from diseases. Mothers FGD, September 2017, intervention village</td>
</tr>
</tbody>
</table>

### Human resources

<table>
<thead>
<tr>
<th>Author</th>
<th>Year</th>
<th>Topic</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gurley</td>
<td>2020</td>
<td>Gurley 2020 Staff</td>
<td>In addition to suboptimal ANM clinical quality, nine ANM posts were vacant during the study period, resulting in suboptimal availability of services in some communities, which was beyond our influence</td>
</tr>
<tr>
<td>Olken</td>
<td>2014</td>
<td>Rahayu 2008 Absenteeism</td>
<td>the community is often disappointed with the services of puskesmas officers who do not attend posyandu services as this means that immunisations must be postponed until the posyandu session scheduled for the next month.</td>
</tr>
</tbody>
</table>
### Infrastructure and supplies

<table>
<thead>
<tr>
<th>Authors</th>
<th>Year</th>
<th>Theme</th>
<th>Quotes</th>
</tr>
</thead>
<tbody>
<tr>
<td>Okeke et al 2017</td>
<td>Okeke et al 2017</td>
<td>Building infrastructure</td>
<td>That is supposed to be our office. If you get there now, half of it is just sand; even the doors to this clinic are not closing, they are eating up by termite. The condition of two of the facilities was perceived to be detrimental to their ability to deliver care.</td>
</tr>
<tr>
<td>Findley et al 2013</td>
<td>Doctor et al 2011</td>
<td>Vaccine stockout and long wait times</td>
<td>67% of parents were unable to receive all immunisations reported lack of vaccine as a problem, and 13% had difficulties with the long wait.</td>
</tr>
</tbody>
</table>

### Multiple barriers

<table>
<thead>
<tr>
<th>Authors</th>
<th>Year</th>
<th>Theme</th>
<th>Quotes</th>
</tr>
</thead>
<tbody>
<tr>
<td>Okeke et al 2017</td>
<td>Okeke et al 2017</td>
<td>Cost of vaccination and logistics</td>
<td>costs associated with use of services and lack of adequate transportation, hindered uptake.</td>
</tr>
<tr>
<td>Arifeen 2019</td>
<td>Billah 2018</td>
<td>Many simultaneous barriers</td>
<td>Furthermore, the lack of trust in formal health care providers, high out of pocket costs of health care in the public sector, and challenges with accessibility has led to the rapid proliferation of informal health providers to fill the gap between supply and demand across both rural and urban areas</td>
</tr>
<tr>
<td>Tandon 1988</td>
<td>Islam 2013</td>
<td>Many simultaneous barriers</td>
<td>low capacity to supervise monitor and implement micro plans at district level, lack of effective vaccine distribution to immunisation sites, ageing and poorly maintained Cold Chain, lack of adequately trained human resources, low managerial and support capacity at the state and district immunisation units and weak management of fund flows.</td>
</tr>
</tbody>
</table>

Supplementary Table 6: Characteristic quotes of themes identified related to barriers to immunisation for interventions that used community engagement in their design.
<table>
<thead>
<tr>
<th>Primary paper</th>
<th>Paper from which citation comes</th>
<th>Theme demonstrated</th>
<th>Quote</th>
<th>Page number</th>
</tr>
</thead>
<tbody>
<tr>
<td>Financial</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Anderson 2009</td>
<td>Andersson 2009</td>
<td>Other costs</td>
<td>Discussions in our focus groups confirmed the importance of poverty as a barrier to vaccination in many cases, as parents described being unable to afford the costs of the supposedly “free” immunisations: travel costs, opportunity costs, and demands for unofficial payments.</td>
<td>10</td>
</tr>
<tr>
<td>Fear</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Dipeolu 2017</td>
<td>Oladepe 2019</td>
<td>Fear of side effects</td>
<td>Prominent factors perceived as affecting timely and full completion of routine immunisations for children aged below 12 months include lack of awareness of immunisation (61.6%) and of subsequent doses (58.4%), and fear of side effects (59.7%). Others include rumours (for example, an association with future infertility)</td>
<td>464</td>
</tr>
<tr>
<td>Adamu 2019</td>
<td>Adamu Dissertation</td>
<td>Past experience</td>
<td>Experiences with the side effects of vaccines can influence the behaviour of caregivers and reduce their motivation to immunize eligible children.</td>
<td>134</td>
</tr>
<tr>
<td>Logistics</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Andersson 2009</td>
<td>Mitchell 2009</td>
<td>Physical access to facilities</td>
<td>In both urban and rural areas, access to a government facility providing vaccinations, a key equity factor, was a determining factor for uptake.</td>
<td>8</td>
</tr>
<tr>
<td>Nagar 2018</td>
<td>Nagar 2016</td>
<td>Convenience</td>
<td>not knowing where to go, not having time or mutually convenient time, facing long wait times at the camp, having fear of side effects, and acting under misguided advice</td>
<td>5</td>
</tr>
<tr>
<td>Human resources</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

188
Staff shortages are common, especially for lady health visitors and pharmacists. Few of the PHCs surveyed are fully staffed.

**Infrastructure or supplies**

<table>
<thead>
<tr>
<th>Author</th>
<th>Paper from which citation comes</th>
<th>Theme demonstrated</th>
<th>Quote</th>
<th>Page number</th>
</tr>
</thead>
<tbody>
<tr>
<td>Domek 2019</td>
<td>Domek et al 2018</td>
<td>Stock outs</td>
<td>Of note, Guatemala experienced significant political instability during our study period, which led to considerable vaccine shortages experienced by all of our clinics.</td>
<td>3</td>
</tr>
</tbody>
</table>

**Poor quality**

<table>
<thead>
<tr>
<th>Author</th>
<th>Paper from which citation comes</th>
<th>Theme demonstrated</th>
<th>Quote</th>
<th>Page number</th>
</tr>
</thead>
<tbody>
<tr>
<td>Dipeolu 2017</td>
<td>Dipeolu 2017</td>
<td>Attitudes of health workers</td>
<td>Attitudes and behaviours of healthcare workers such as treating mothers in an unfriendly, disrespectful, or even abusive manner are frequently cited as discouraging children’s vaccination. Healthcare workers reportedly screamed at mothers who forgot the child’s card, missed a scheduled vaccination</td>
<td>49</td>
</tr>
<tr>
<td>Gurely 2020</td>
<td>Gurley 2020</td>
<td>Quality affected attendance</td>
<td>For example, the existence of ANM vaccinators did not ensure their quality, and we observed that suboptimal clinical ANM quality discouraged retention across the vaccine schedule and perhaps negatively impacted beneficiaries’ trust in the health system.</td>
<td>37</td>
</tr>
</tbody>
</table>

Supplementary Table 7: Characteristic quotes of themes identified related to barriers to immunisation for interventions that used community engagement in their implementation

<table>
<thead>
<tr>
<th>Primary paper</th>
<th>Paper from which citation comes</th>
<th>Theme demonstrated</th>
<th>Quote</th>
<th>Page number</th>
</tr>
</thead>
<tbody>
<tr>
<td>Financial</td>
<td></td>
<td></td>
<td>costs associated with use of services and lack of adequate transportation, hindered uptake.</td>
<td>54</td>
</tr>
<tr>
<td>Author(s)</td>
<td>Year</td>
<td>Source</td>
<td>Description</td>
<td></td>
</tr>
<tr>
<td>----------</td>
<td>------</td>
<td>--------</td>
<td>-------------</td>
<td></td>
</tr>
<tr>
<td>Adamu 2019</td>
<td>Adamu 2019</td>
<td>Other costs associated with vaccination</td>
<td>However, the cost of treating vaccine reactions like fever that might occur following immunisation are borne for the caregivers, which can result in out-of-pocket expenditure.</td>
<td></td>
</tr>
<tr>
<td>Social norms</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Adamu 2019</td>
<td>Adamu 2019</td>
<td>Spousal consent</td>
<td>In addition, we found that the high level of social control that men have over women in this area could also cause MOV among children. Caregivers reported that without their husband’s consent, they still cannot vaccinate their children, even if they’re in a health facility for other preventive or curative services.</td>
<td></td>
</tr>
<tr>
<td>Olken 2014</td>
<td>Rahayu 2008</td>
<td>Religious beliefs</td>
<td>The community trusts and believes in the choice that their parents made to use the dukun beranak, to the point where it becomes the norm.</td>
<td></td>
</tr>
<tr>
<td>Logistics or distance</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Webster 2019</td>
<td>Webster et al 2019</td>
<td>Convenience</td>
<td>Key reasons cited for non-immunisation were related to lack of convenience; the aforementioned distance between homes and health facilities, moving to the fields during farming seasons,</td>
<td></td>
</tr>
<tr>
<td>Alhassan 2019</td>
<td>Alhassan 2019</td>
<td>Transport</td>
<td>In Ghana, access to maternal and child health services is impeded by longer travel times to health facilities.</td>
<td></td>
</tr>
<tr>
<td>Human resources</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Olken 2014</td>
<td>February 2011</td>
<td>Post abandonment</td>
<td>Poor facilities adversely affected the availability and utilization of MCH and basic education services. The midwife did not want to live in the village because here there is no electricity and no water (FGD Female-NTT).</td>
<td></td>
</tr>
<tr>
<td>Source</td>
<td>Source</td>
<td>Facilitator</td>
<td>Description</td>
<td></td>
</tr>
<tr>
<td>--------</td>
<td>--------</td>
<td>-------------</td>
<td>-------------</td>
<td></td>
</tr>
<tr>
<td>Herrara-Almanza &amp; Rossales-Rueda 2018</td>
<td>Herrara-Almanza &amp; Rossales-Rueda 2018</td>
<td>Poor transportation infrastructure and a shortage of medical personnel limit the basic health access to the population living in remote areas</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Infrastructure or supplies**

| Webster 2019 | Webster et al 2019 | Vaccine stockouts | Furthermore, vaccine stock-outs at health facilities discouraged caregivers from attending vaccination services. |

**Multiple barriers**

| Webster 2019 | Webster et al 2019 | Demand and supply side | Health system factors affecting immunisation service delivery and uptake in northern Uganda include interruptions in the vaccine cold chain due to poor management of equipment and supplies, limited supervision of health teams by the district health teams, low staffing levels, long distances to the health facilities especially in the hard-to-reach areas, and limited resources to support outreach services. There is also a lack of good quality data to support decision-making. |
| Olken 2014 | Rahayu 2008 | Poor quality and post abandonment | Dissatisfaction is often related to the midwife’s character, ineffective medicine, minimal experience of the midwife, difficulties in reaching the midwife, and a midwife’s absence from the post. |
| Herrara-Almanza & Rossales-Rueda 2018 | Herrara-Almanza & Rossales-Rueda 2018 | Transport and personnel | Poor transportation infrastructure and a shortage of medical personnel limit the basic health access to the population living in remote areas |

**Facilitators of immunisation**

Fewer studies discussed facilitators of immunisation as compared to barriers. Much of the information coded came from studies in which the intervention was engagement. To a substantial extent, the sources of this information were key individual interviews and focus group discussions. These were often supported by surveyor or enumerator’s observations. In some instances, discussion on facilitators in the primary papers was informed by existing literature.
A sensitivity analysis, comprising 19 qualitative papers with a risk of bias assessment score of 20 or higher, indicated similar results. Therefore, we view the evidence to be of high strength. Facilitators related to behavioural, social and practical factors faced by caregivers were more common than the facilitators related to the delivery of immunisation services (Supplementary Figure 86). A segregated analysis of the facilitators by the engagement type was not possible due to the limited information coded on this theme.

**Behavioural, social and practical barriers faced by caregivers**

The most consistent facilitator across all intervention types was the caregiver’s *understanding of importance of immunisation*. In general, studies associated caregivers’ awareness and perception of the benefits of vaccination to an enabling environment for immunisation uptake (Adamu 2019; Banerjee 2020). In most instances caregivers were able to link benefits of vaccines to the diseases it prevented (Anderson (Cockroft 2009); Pramanik 2020).

*Lack of logistics or distance related issues* and enabling *social norms* were also consistently sighted as facilitators of immunisation. When it came to logistics, easy access either in the form of proximity to a health care facility (Anderson (Cockroft 2009)) or provision of transportation to an immunisation camp (Olken 2014; Rahayu 2008) were cited as enabling factors. *Social norms* were mentioned as a facilitator in instances where mothers had a significant influence of over decisions regarding her child’s health. In addition, peer influence was also mentioned as a factor (Banerjee 2020). Though less consistently cited than the other factors, *favourable socio-economic characteristics* such as maternal education (Admassie 2009) were also cited as facilitators of immunisation.

**Constraints in delivery of immunisation services**

Delivery of immunisation services was less consistently reported as a facilitator as compared to behavioural, social and practical factors. The most consistent delivery related facilitator was the *availability of infrastructure or supplies* which included provision of good quality of services, availability of adequate healthcare providers at the point of care and availability of vaccine supplies.
Supplementary Figure 86: Summary figure: Facilitators related to behavioural, social and practical factors faced by caregivers were more common than those related to delivery of immunisation services

The size of the rectangles reflects the number of times a unique theme was mentioned by an IE and its associated papers. Larger rectangles indicate that a theme was mentioned more times. Nested boxes reflect the proportion of times the main theme text also fell into the sub-them.
### Facilitators of immunisation by type of community engagement

<table>
<thead>
<tr>
<th>Failure</th>
<th>Intervention features</th>
<th>Implementation or scale-up challenges</th>
</tr>
</thead>
<tbody>
<tr>
<td>Social norms: Not accounting for existing constraints or uncontrollable contextual trends</td>
<td>Duration, frequency, ...</td>
<td>Other delays, disruptions or im...</td>
</tr>
<tr>
<td>Health system issues: Poor quality of services</td>
<td>Design - other</td>
<td></td>
</tr>
<tr>
<td>Migration trends: Health ...</td>
<td>Payment delays</td>
<td>Competing pr...</td>
</tr>
<tr>
<td>Civil unrest or political instability</td>
<td>Caregiver competence</td>
<td></td>
</tr>
<tr>
<td>Fear of AEFI</td>
<td>Migration ...</td>
<td></td>
</tr>
<tr>
<td>High baseline coverage: Lack of awareness ...</td>
<td>Resource constraint</td>
<td></td>
</tr>
</tbody>
</table>

**Supplementary Figure 87: Engagement as intervention**

<table>
<thead>
<tr>
<th>Failure</th>
<th>Intervention features</th>
<th>Study design</th>
</tr>
</thead>
<tbody>
<tr>
<td>Social norms: Not accounting for existing constraints or uncontrollable contextual trends</td>
<td>Duration, frequency a...</td>
<td>Other methodology ...</td>
</tr>
<tr>
<td>Health system issues: Poor quality of services</td>
<td>Design - other</td>
<td></td>
</tr>
<tr>
<td>Migration trends: Health service access ...</td>
<td>Payment delays</td>
<td></td>
</tr>
<tr>
<td>Civil unrest or political instability</td>
<td>Caregiver competence</td>
<td></td>
</tr>
<tr>
<td>Fear of AEFI</td>
<td>Migration ...</td>
<td></td>
</tr>
<tr>
<td>High baseline coverage: Lack of awareness ...</td>
<td>Resource constraint</td>
<td></td>
</tr>
</tbody>
</table>

**Supplementary Figure 88: Engagement in design**
### Facilitators of immunisation

<table>
<thead>
<tr>
<th>Behavioural, social and practical factors faced by caregivers</th>
<th>Delivery of immunisation services</th>
</tr>
</thead>
<tbody>
<tr>
<td>Lack of financial...</td>
<td>Availability of services</td>
</tr>
<tr>
<td>Favourable popula...</td>
<td>Availability of infrastr...</td>
</tr>
<tr>
<td>Understanding of importance</td>
<td></td>
</tr>
<tr>
<td>Lack of logistics or distance related challenges</td>
<td></td>
</tr>
<tr>
<td>Social norms</td>
<td></td>
</tr>
<tr>
<td>Knowledge of I...</td>
<td></td>
</tr>
</tbody>
</table>

**Supplementary Figure 89: Engagement in implementation autonomy**

**Supplementary Table 8: Characteristic quotes of themes identified related to facilitators of immunisation for community engagement interventions**

<table>
<thead>
<tr>
<th>Primary paper</th>
<th>Paper from which citation comes</th>
<th>Theme demonstrated</th>
<th>Quote</th>
<th>Page number</th>
</tr>
</thead>
<tbody>
<tr>
<td>Adamu et al. 2019</td>
<td>Adamu Dissertation 2019</td>
<td>Awareness of benefits</td>
<td>In addition, caregiver beliefs about the capabilities of immunization was overwhelmingly strong: “… this vaccination is very important because it prevents infection from measles, cough, hepatitis, fever, pneumonia and yellow fever.</td>
<td>129</td>
</tr>
<tr>
<td>Oyo-Ita et al. 2020</td>
<td>Oyo-Ita et al. 2020</td>
<td>Positive perception of vaccination</td>
<td>It was found from the baseline qualitative study that respondents were generally knowledgeable about and had a positive attitude towards vaccination. They believed vaccines prevented their children from acquiring deadly infections and attributed low numbers of deaths of children to vaccination</td>
<td>31</td>
</tr>
</tbody>
</table>

**Socio-economic characteristics**
Andersson et al. Cockroft 2009  
2009  
Importance of maternal education  
Mother’s education was related to measles vaccination in all four districts, in urban and rural sites.

Social norms

Pramanik 2020 Pramanik 2020  
Household decision-making  
Key informants further mentioned that in rare cases, it was seen that presence of a knowledgeable daughter in law in the family compelled families to vaccinate their children.

Availability of infrastructure, supplies or services

Olken et al. 2014 Rahayu 2008  
Availability of medical supplies  
The midwife has comprehensive equipment. She has infusion equipment, blood pressure monitor, injections, medicines, scales for infants. The midwife also provides a room in the polindes for mothers who have just given birth. (Women’s FGD, Kuanek, East Miomaffo, TTU, NTT)

Andersson et al. Cockroft 2009  
2009  
Outreach by healthcare workers  
in rural sites a vaccination team visiting the community also increased the likelihood that the child had received measles vaccine; this effect was much stronger in rural Khanewal.

Adamu et al. 2019 Adamu 2019 Dissertation  
Availability of vaccination services  
In this study, some caregivers indicated that recommended birth doses of vaccines are provided in the labor room as soon as the child is born. This practice saves time and reduces the chance of missed opportunities in this service delivery point.
Reasons for intervention success

Most studies with significant positive quantitative impacts provide some plausible explanation for why an intervention succeeded. These were often based on authors’ notes, experiences and impressions of why their hypothesis was proven true. We expect the authors to be a good source of this information because of their knowledge of causal linkages and assumptions made in the intervention theory of change. Individual interviews and focus group discussions supported the authors’ conclusions. The sensitivity analysis, which included qualitative papers with an assessment of score of 20 or higher, showed similar results. Given this, we consider the evidence to be of high strength.

Overall, reasons for intervention success were consistently attributed to specific intervention features across all engagement types. These included community engagement, leadership and supportive supervision, customisation to local context, incentives, health system integration and other design features. Though not as common, success was also attributed to existing or changing favourable characteristics within a given context such as availability of and access to good quality health services as well as positive participant views with respect to the intervention (Supplementary Figure 90; Supplementary Table 9).

<table>
<thead>
<tr>
<th>Success</th>
<th>Intervention features</th>
<th>Other design features</th>
<th>Incentives</th>
<th>Health system</th>
<th>Existing or changing favourable factors or care...</th>
</tr>
</thead>
<tbody>
<tr>
<td>Community engagement</td>
<td>Community...</td>
<td>Comm...</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>New cadres</td>
<td>Needs assessment...</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Leadership and supportive supervision</td>
<td>Customisation to local context</td>
<td>Behaviour...</td>
<td>Health...</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Health worker training</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Supplementary Figure 90: Reasons for success: a summary figure**

The size of the rectangles reflects the number of times a unique theme was mentioned by an IE and its associated papers. Larger rectangles indicate that a theme was mentioned more times. Nested boxes reflect the proportion of times the main theme text also fell into the sub-theme.

**Supplementary Table 9: Comparison of the reasons for intervention success by type of community engagement.**
<table>
<thead>
<tr>
<th><strong>Intervention features</strong> – community engagement</th>
<th>Community engagement as the intervention</th>
<th>Community engagement in design</th>
<th>Community engagement in implementation autonomy of interventions</th>
</tr>
</thead>
<tbody>
<tr>
<td>Community dialogues</td>
<td>Community dialogues Creation of new cadres</td>
<td>Community dialogues Needs assessments</td>
<td>Community involvement in planning &amp; implementation</td>
</tr>
<tr>
<td>Leadership &amp; supportive supervision</td>
<td>Encouragement from leaders</td>
<td>Supportive supervision</td>
<td>Involvement of leadership Health worker autonomy</td>
</tr>
<tr>
<td>Health system integration</td>
<td>Involvement of government in implementation</td>
<td>Separate workforce dedicated to the intervention</td>
<td>Intersectoral integration of health programmes</td>
</tr>
<tr>
<td>Health worker training</td>
<td>Training on monitoring health cards of children</td>
<td>Training of all health personnel involved in the intervention</td>
<td>Training of all health personnel involved in the intervention</td>
</tr>
<tr>
<td>Health system enablers – Favourable contextual factors</td>
<td>Good quality of services Health service access</td>
<td>Good quality of services</td>
<td>Good quality of services</td>
</tr>
<tr>
<td>Socio-economic factors</td>
<td>Female literacy</td>
<td>General literacy levels High maternal education</td>
<td>Primary schooling of girls and women</td>
</tr>
<tr>
<td>Improvements in intervention Implementation over time</td>
<td>-</td>
<td>Health worker performance improvement over time</td>
<td>Resolution of fidelity/adherence issues</td>
</tr>
<tr>
<td>Positive participant/beneficiary views</td>
<td>Caregiver perception</td>
<td>Health worker perception</td>
<td>Health worker perception</td>
</tr>
</tbody>
</table>
Community engagement as the intervention

Studies in which engagement was the intervention consistently attributed success to characteristics associated with intervention features and relatively less consistently to existing or changing favourable characteristics within a given context.

Notably, success was attributed to intervention features associated with community engagement such as community dialogues which were participatory in nature and improved relationship between the health system and communities (Assegai 2018 (Padayachee 2013); Findley 2013). Involvement of community members such as traditional and religious leaders in planning and implementation of an intervention also came up as a reason for success (Oyo-Iua 2020). In a few instances, success was also attributed to acceptance of health worker cadres by the communities as long as these health workers belonged to the communities they served (Biemba 2016).

Studies also attributed intervention success to non-engagement intervention features including incentives given to caregivers, leadership and supportive supervision which improved overall health service delivery and health worker performance. As noted above, some studies also cited existing or changing favourable characteristics within a given context as reasons for success. Among these, those related to the health system, particularly, good service quality (Rao 2016) and, to a lesser extent, access to services (Banwat 2015 (Banwat 2014)) were consistently reported. Interestingly, none of the studies in this engagement category mentioned implementation improvements as a reason for intervention success.

Supplementary Figure 91: Reasons for success of interventions with engagement as intervention

Supplementary Table 10: Characteristic quotes of themes identified related to reasons for success for interventions were community engagement.
intervention features

<table>
<thead>
<tr>
<th>Authors</th>
<th>Year</th>
<th>Study Title</th>
<th>Study Design</th>
<th>Findings</th>
</tr>
</thead>
<tbody>
<tr>
<td>Assegai</td>
<td>2018</td>
<td>Community dialogue</td>
<td>Interviewees spoke of the tangible benefits of the community dialogues and proactive household approaches as expanding access, improving relationships with communities, and increasing knowledge and uptake of services.</td>
<td></td>
</tr>
<tr>
<td>Oyo-Ita</td>
<td>2020</td>
<td>Community involvement in planning and implementation</td>
<td>Inclusion of TRLs in the planning, implementation and evaluation of an intervention is useful in ensuring support from the community.</td>
<td></td>
</tr>
<tr>
<td>Biemba</td>
<td>2016</td>
<td>Leadership and supportive supervision</td>
<td>CHAs reported that supervision by experienced health workers was valuable to them because it reinforced skills they learned in training and provided general encouragement. This study finds that the quality and frequency of supervision is heavily influenced by the proximity of supervisors’ work station to that of the CHAs.</td>
<td></td>
</tr>
<tr>
<td>Banerjee</td>
<td>2020</td>
<td>Incentives</td>
<td>Most primary caregivers expressed positive opinions about recharges. Some admitted to finding them motivating or influencing their decision to immunise their child (whether or not they actually received them). This was linked to generally liking free things (‘What’s not to like?’), their self-assessed poverty level, or their trust in the government ‘investing in them’.</td>
<td></td>
</tr>
</tbody>
</table>
Existing or changing favourable factors or contextual trends

<table>
<thead>
<tr>
<th>Author(s)</th>
<th>Year</th>
<th>Good quality of services and infrastructure</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Rao et al.</td>
<td>2014</td>
<td></td>
<td>Moreover, the ASHA worker is envisioned not to work in isolation but as an integral part of the public health system, working together with other health personnel to both ease their work-pressures and assist them in performing their duties more efficiently. For instance, the ASHA worker works with the ANM and the anganwadi worker to organize monthly village health days. Therefore, the extent to which the increase in information provision via other health workers is attributable to the ASHA worker is not quantifiable in this context.</td>
</tr>
<tr>
<td>Findley et al.</td>
<td>2013</td>
<td></td>
<td>Simultaneous improvement in the quality of care provided by the CHW and nurse-midwives at the health post gave women the confidence that they could go to the health post to seek advice and care.</td>
</tr>
<tr>
<td>Banwat</td>
<td>2015</td>
<td>Good quality of services and infrastructure</td>
<td>A child was also found to be most likely to be fully and timely immunized if he/she lived within 30 minutes walking distance from the health facility.</td>
</tr>
</tbody>
</table>

Community engagement in design

The most consistent reasons for success were related to intervention features with the most reported characteristics being community engagement and provision of incentives to caregivers. Within community engagement, participatory community dialogues (Andersson 2009), stakeholder consultations on intervention design (Modi 2019) and community involvement in planning and implementation (Adamu 2019) were associated with intervention success. Receiving incentives was perceived to be motivating and empowering by the intended beneficiaries (Banerjee 2020; Nagar 2020).
Besides community engagement and incentives, other intervention features that were consistently reported included leadership and supportive supervision, health worker training and customisation to local context.

Very few studies discussed or attributed success to reasons other than intervention features. Nevertheless, the few studies that did, referred to existing or changing favourable contextual factors such as a high prevalence of maternal education (Johri 2020) or improved implementation by resolution of issues related to adherence (Modi 2019). A few studies also discussed positive participant or beneficiary views that could have potentially contributed to intervention success.

Supplementary Figure 92: Reasons for success of interventions with engagement in design

Supplementary Table 11: Characteristic quotes of themes identified related to reasons for success for interventions that used community engagement in their design.

<table>
<thead>
<tr>
<th>Primary paper</th>
<th>Paper which comes from</th>
<th>Sub-theme demonstrated</th>
<th>Quote</th>
</tr>
</thead>
<tbody>
<tr>
<td>Andersson 2009</td>
<td>-</td>
<td>Community dialogue</td>
<td>The structured discussion rounds sometimes led to 8 action plans in the intervention communities beyond stimulating discussion about vaccinations within households. Particularly in those villages with poor access to vaccination services, plans included sharing transport to vaccination points and providing care for some children while parents took others to be vaccinated. These community initiatives may have helped to maintain vaccination levels in the face of generally falling levels.</td>
</tr>
</tbody>
</table>
Participants expressed satisfaction with the supervisory plan that was put in place for the quality improvement program as it enabled quick feedback. Supervisory visits were conducted by different stakeholders that are higher-ranking officials within the health systems. These include local government and zonal primary health-care management board officials. “We receive supervision from local government, they used to come and supervised us to check how we conduct our duties.”– PHC 3

Referrals to clinics by the teams were regarded as an important aspect of bringing services closer to the communities. Respondents said that they would not have gone to the clinic had they not been referred by the WBOT, for example for deworming of children, or continuation of care for complicated cases.

By and large the main, explicitly stated driver for primary caregivers to get their child vaccinated is their positive perception or attitude toward immunization: that immunization will benefit their child. This acts as a push even for primary caregivers who have limited knowledge, who don’t understand exactly how or why immunization is beneficial. Some of these primary caregivers hold strong personal conviction of the importance of immunization, despite their lack of knowledge, while others are undecided or unsure about the importance of immunization, but are still overall positive. In both cases, it seems that the positive perception or attitude is influenced by others: what others say and do. This is often mediated by trust: either trust in government and authority (e.g. ANM), household or family members, other primary caregivers in the community, or people in positions of power in the community (e.g. religious leaders in Mewat). Interestingly, improved knowledge and awareness was stated by a number of respondents as a key factor to convince primary caregivers (themselves, or others) to attend immunization camps. This demonstrates, as was mentioned in response to ANM performance, that
Similar to the previous two engagement categories, reasons for success in interventions with engagement in implementation were also consistently attributed to certain intervention features. These characteristics were associated with certain aspects of community engagement itself such as conducting stakeholder consultations (Modi 2019) or involvement of community members such as health workers in intervention planning and implementation (Adamu 2019). Other aspects of intervention features such as integration with the health system, customisation to the local context and leadership and supportive supervision were also consistently cited.

Like the engagement in design category, crediting intervention success to existing or changing favourable contextual factors or improved implementation was less consistently reported among studies in which community engagement was involved in the implementation of the intervention. Nevertheless, health system enablers such as good quality of services and access to health services were mentioned as factors which could have contributed to intervention success.

Community engagement in implementation autonomy of intervention

Similar to the previous two engagement categories, reasons for success in interventions with engagement in implementation were also consistently attributed to certain intervention features. These characteristics were associated with certain aspects of community engagement itself such as conducting stakeholder consultations (Modi 2019) or involvement of community members such as health workers in intervention planning and implementation (Adamu 2019). Other aspects of intervention features such as integration with the health system, customisation to the local context and leadership and supportive supervision were also consistently cited.

Like the engagement in design category, crediting intervention success to existing or changing favourable contextual factors or improved implementation was less consistently reported among studies in which community engagement was involved in the implementation of the intervention. Nevertheless, health system enablers such as good quality of services and access to health services were mentioned as factors which could have contributed to intervention success.

Implementation improvements

Gibson et al. 2017 - Socio-economic factors

The success of SMS reminders to elicit a behaviour is multifactorial; the content of the message, the type of behaviour being reminded, indirect and direct costs incurred, literacy level, and other contextual factors all being potential explanatory factors.

Modi 2019 - Operational improvements

Finally, we identified critical operational requirements that improved adherence to the intervention; this included supportive supervision, timely resolution of technology problems, and change management, including monetary and/ or nonmonetary incentives depending on the context.
Supplementary Table 12: Characteristic quotes of themes identified related to reasons for success for interventions that used community engagement in implementation.

<table>
<thead>
<tr>
<th>Primary paper</th>
<th>Paper from which citation comes</th>
<th>Sub-theme demonstrated</th>
<th>Quote</th>
<th>Page number</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intervention features</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Adamu 2019</td>
<td>-</td>
<td>Community involvement in planning and implementation</td>
<td>Stakeholders were systematically involved in the planning and execution of the QI program. The change ideas were selected by frontline health workers and they tailored them to their local context. Also, change ideas are multi-faceted and multimodal, and this is suitable for addressing complex problems in complex health systems.</td>
<td>9</td>
</tr>
<tr>
<td>Sankar 2013</td>
<td>-</td>
<td>Needs assessments, pilots or stakeholder consultations</td>
<td>The project interventions were developed in keeping with the especial needs of the population, and rooted in the conditions and circumstances of the area. It drew heavily from the understanding and opinions of local people. The results reflect the fruits of a highly contextualised program.</td>
<td>40</td>
</tr>
</tbody>
</table>
District Health Offices and DHCCs were less well established compared to the Provincial Public Health Directorates and PPHCCs; the ministry did not have adequate resources to equip them well. Despite these challenges, the leadership and involvement of the ministry in the intervention mattered. The provincial and district health governance leaders were inspired to improve their governance because the ministry leaders were interested in the pilot intervention.

Existing or changing favourable factors or contextual trends

Admassie 2009 - Good quality services and infrastructure

The impact of the programme on the proportion of children who got vaccines against major childhood illnesses also varies according to supply side variables (see Table 6). The programme has larger effects in villages with one HEW and with better quality health posts.

Herrera-Almanza and Rosales-Rueda 2018 - Health service access and availability

However, the fact that we find results on vaccination uptake is an indication that the health workers were somehow present in remote areas.

Admassie 2009 - Socio-economic factors

This implies that encouraging and supporting primary schooling for girls and women will enhance the programme’s impact on proportion of children vaccinated against major childhood illnesses.

Implementation improvements

Mayumana 2017 - Binyaruka 2017 - Managing vaccine stockouts

The verification system under P4P also meant that district supervision was intensified, providing more opportunities for district managers to identify and address stock-outs of a wider range of drugs.
Reasons for intervention failure

Most studies that did not find a significant positive impact attributed the failure of the intervention to a specific cause. The most consistently reported reason for failure across all engagement types was *not accounting for contextual constraints*, though the exact nature of these reasons varied. Social norms such as misconception regarding immunisation and health worker attrition were reported as common challenges in interventions that were community engagement. Interventions that used engagement in design consistently reported poor access to health services and caregivers’ skepticism regarding benefits of immunisation as the reasons for failure. Poor quality of health services was noted as an issue in interventions where engagement occurred in implementation.

The other consistently reported reasons for failure were attributed to inadequate *intervention features* and *implementation challenges*. Among the intervention features, inadequate duration, frequency or exposure to the intervention were the most notable reasons for failure across all engagement types. Though relatively less common, the nature of the community engagement itself was also attributed to failure in interventions that used engagement in implementation or as intervention. Among implementation challenges, disruption due to inadequate implementation instructions, difficulty in accessing the implementation sites due to geographic proximity and competing priorities of health workers were noted as some of the prevalent challenges (Supplementary Figure 94; Supplementary Table 13).

Another less consistently reported reason for failure was study design issues such as contamination between treatment and control other methodological challenges such as biased sampling frame and issues with enrolling study populations. The reasons for failure were largely based on authors’ notes or experiences and individual interviews. We expect the authors to be a good source of this information because of their knowledge of causal linkages and assumptions made in the intervention theory of change. Focus group discussions and implementers’ experiences supported these findings. The sensitivity analysis, which included qualitative papers with an assessment of score of 20 or higher, showed similar results. Except for one sub-theme which was dominated by one or two studies, other sub-themes were consistently distributed across the evidence base (indicated in Supplementary Table 13). Therefore, we consider the evidence to be of high strength.

---

**Supplementary Figure 94: Summary figure: reasons for failure**

**Supplementary Table 13: Comparison of the reasons for intervention failure by type of community engagement.**
<table>
<thead>
<tr>
<th></th>
<th>Community engagement as the intervention</th>
<th>Community engagement in design</th>
<th>Community engagement in implementation autonomy of the intervention</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Not accounting for contextual constraints – health system issues</strong></td>
<td>Health worker attrition</td>
<td>Issues with accessing health services</td>
<td>Poor quality of services</td>
</tr>
<tr>
<td><strong>Not accounting for contextual constraints – prevalent social norms</strong></td>
<td>Fear of sterilization</td>
<td>Skepticism towards perceived value of immunisation</td>
<td>-</td>
</tr>
<tr>
<td><strong>Not accounting for contextual constraints – baseline immunisation coverage</strong></td>
<td>High baseline coverage</td>
<td>High baseline coverage</td>
<td>High BCG coverage</td>
</tr>
<tr>
<td><strong>Not accounting for contextual constraints – unrest, insecurity or other</strong></td>
<td>Civil unrest Natural calamity</td>
<td>Political instability</td>
<td>-</td>
</tr>
<tr>
<td><strong>Intervention features – Duration, dose or frequency of intervention</strong></td>
<td>Inadequate intervention duration or exposure</td>
<td>Inadequate intervention frequency</td>
<td>Inadequate intervention duration or exposure</td>
</tr>
<tr>
<td><strong>Intervention feature – issue in design of engagement</strong></td>
<td>Lack of collective action</td>
<td>-</td>
<td>Elite capture</td>
</tr>
<tr>
<td><strong>Implementation challenges – disruptions or delays</strong></td>
<td>Resource constraint due to multiple programmes running at the same time</td>
<td>Fidelity issues</td>
<td>Staff turnover*</td>
</tr>
<tr>
<td><strong>Study design –</strong></td>
<td>Contamination due to similar programmes</td>
<td>Contamination due to similar programmes</td>
<td>Contamination between treatment and control</td>
</tr>
</tbody>
</table>
Contamination

*Most of the references coded under this theme were from Olken 2014 and associated papers.

**Community engagement as the intervention**

_Not accounting for contextual trends_ were the most consistently reported reason for failure. Social norms, health system issues, migration, civil unrest and fear of adverse events following immunisation were reported consistently within existing constraints. Most of the challenges related to social norms mentioned misconception about immunisation (Banerjee 2020) and limited decision-making powers of mothers (Gurley 2020) as the reason for low immunisation uptake. Civil unrest, political instability and natural calamities also hindered a few interventions from achieving intended impacts (Pramanik 2020, Morris 2004).

The other reasons for failure comprised issues related to _intervention features, implementation issues_ and, to a lesser extent, study _design_. Issues with intervention feature ranged from inadequate duration or frequency of the intervention to certain aspects of community engagement failing to elicit a positive response among study participants (Pramanik 2020; More 2012). Implementation issues were consistently reported as reasons for failure. Some of the challenges mentioned were competing demands by different programmes running simultaneously (Carnell 2014), high staff turnover (Gurley 2020; Olken 2014) and low implementation fidelity (Oyo-Ita 2020). These issues are also discussed in detail under the ‘uptake and fidelity challenges theme’.

---

**Supplementary Figure 95: Reasons for failure of interventions with engagement as intervention**

Supplementray Table 14: Characteristic quotes of themes identified related to reasons for failure for interventions were community engagement.

---

209
<table>
<thead>
<tr>
<th>Primary paper</th>
<th>Paper from which citation comes</th>
<th>Sub-theme demonstrated</th>
<th>Quote</th>
<th>Page number</th>
</tr>
</thead>
<tbody>
<tr>
<td>Carnell 2014</td>
<td>-</td>
<td>Attrition of personnel</td>
<td>The ESHE and RHB teams both faced personnel challenges in Amhara, which may have affected management and performance.</td>
<td>559</td>
</tr>
<tr>
<td>Gurley 2020</td>
<td>-</td>
<td>Infrastructure constraints</td>
<td>Community constraints around screening spaces and economic activities also introduced challenges.</td>
<td>73</td>
</tr>
</tbody>
</table>

**Implementation or scale-up challenges**

**Not accounting for existing constraints or uncontrollable trends**

| More 2012     | -                               | Lack of access to health care | The third issue was the complexity of urban health care. Antenatal care was the norm and the nadir for institutional delivery in trial clusters was 75%. Around 57% of antenatal care and 30% of deliveries were in the private sector (this in a slumdwelling population). Open access to private providers, and to institutions at all levels of the public sector hierarchy, is a challenge to systematic health care delivery. Our findings confirmed the tendency to bypass public maternity homes, which should handle uncomplicated deliveries, in favour of tertiary institutions. | E1001257 |
| Pramanik 2020 | -                               | Natural calamity             | Assam has a wonderful collection of festivals that are of the utmost importance. And more tragically when the Brahmaputra River floods the priority is survival. The project simply has to adapt to these ebbs and flows, but there is a pressure to deliver to a schedule that inevitably leads to a box ticking approach that is inconsistent with the development of ownership by the community. | 68 |
Community engagement in design

Most consistently reported reasons for failure of interventions in which communities were engaged in the design broadly fell under not accounting for contextual constraints. Among these, social norms were the most consistently reported reason with vaccine hesitancy being one of the driving forces behind immunisation refusal.

Intervention features

- **Fear of AEFI**
  - Oyo-Ita 2020
  - These responses are indications that fear of side effects can hinder vaccine uptake. This may have contributed to non-impact on the proportion of fully vaccinated children in this study, as the TRLs in the post-intervention qualitative study still mentioned this as a common reason for poor uptake of vaccines.

- **Engagement failure due to elite capture**
  - Olken 2014, Grayman 2013
  - In general, researchers discovered that Generasi provides opportunities for elites to retain and fortify their social standing. As brokers delivering CDD resources into their communities, local elites leverage Generasi to accumulate additional status for themselves. Without exception, every Generasi actor that the researchers met held other leadership roles in their community and maintained close ties with the local structures of village and sub-district governance.

- **Inadequate exposure to the intervention**
  - Robertson 2013
  - Our study was limited by the short intervention period. Whether the effects of the programmes would change with time is unclear. The follow-up survey was done 2 months after interventions had finished, so fear of penalties should not have biased responses from CCT households, although the effects of the programmes could have attenuated by the time of the survey.

Study design

- **Contamination due to other programmes**
  - More 2017
  - although we found no evidence of contamination of control clusters by the intervention, government schemes and the activities of municipal and non-governmental providers might have improved health in control clusters.
(Banerjee 2020; Gurley 2020). Other reasons for failure within this theme were high baseline immunisation coverage, health system issues related to quality of services, and civil unrest or political instability. Prevailing socio-economic trends, lack of awareness, fear of AEFI and caregiver competing priorities were also attributed to intervention failure, though less consistently reported as compared to other factors.

Besides existing constraints, a few studies also attributed failure intervention features with the insufficient duration, frequency and exposure of intervention as a possible reason for failure. A much smaller subset of studies also referred to issues related to study design and implementation or scale-up challenges.

Supplementary Figure 96: Reasons for failure of interventions with engagement in design

Supplementary Table 15: Characteristic quotes of themes identified related to reasons for failure for interventions that used community engagement in their design.

<table>
<thead>
<tr>
<th>Primary paper</th>
<th>Paper from which citation comes</th>
<th>Sub-theme demonstrated</th>
<th>Quote</th>
<th>Page number</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gurley 2020</td>
<td>-</td>
<td>Staff turnover</td>
<td>On the implementer side, staff turnover at NYST and research activities introduced delays in video production.</td>
<td>73</td>
</tr>
<tr>
<td>Engineer 2016</td>
<td>-</td>
<td>Payment delays</td>
<td>Problems with implementation likely dampened any potential effect. The scheme was rolled out in phases, but there were some delays, particularly with the initial payments.</td>
<td>456</td>
</tr>
</tbody>
</table>

212
### Not accounting for existing constraints or uncontrollable trends

<table>
<thead>
<tr>
<th>Author</th>
<th>Year</th>
<th>Field</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gibson 2017</td>
<td></td>
<td>High baseline immunisation coverage</td>
<td>SMS reminders were probably not effective at improving full immunisation and vaccine-specific coverages in this study because of high baseline coverage levels and because SMS reminders might not have addressed the demand side deficiencies in this study area.</td>
</tr>
<tr>
<td>Banerjee 2020</td>
<td></td>
<td>Caregiver competing priorities</td>
<td>Across districts, there are ‘hard-to-convince’ populations who are not affected by incentives. Some ANMs listed specific populations groups they felt continued to be unaffected by the programme, and the incentives specifically, such as migrant workers, daily wage workers and the Muslim community. In Panipat, the top reason given by ANMs for resistance to immunisation by the few still unconvinced (migrants, Muslims) was the fever brought on as an after-effect by a certain vaccine (penta30) and direct costs involved for daily wage workers.</td>
</tr>
<tr>
<td>Gurley 2020</td>
<td></td>
<td>Social norms</td>
<td>It is worth noting that in contrast to the high intent recorded through the household survey, the process evaluation indicated there were individuals who actively chose to not vaccinate their children. This may be a function of social desirability bias in the household survey, though we posit that it may also reflect family constraints or logistical barriers that prevent a mother from taking a vaccine decision regardless of own personal intention to vaccinate. As noted elsewhere, mothers are not the sole or primary decision-makers and other family members reported lower intent to vaccinate.</td>
</tr>
</tbody>
</table>

### Intervention features

<table>
<thead>
<tr>
<th>Author</th>
<th>Year</th>
<th>Field</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Siddiqi 2020</td>
<td></td>
<td>Inadequate exposure to the intervention</td>
<td>Our findings also provide evidence of the strong reliance on the immunization card being the established immunization recall method and we may also postulate that the short duration of the study did not provide enough time to ‘institutionalize’ the use of the bracelets.</td>
</tr>
</tbody>
</table>
Community engagement in implementation autonomy of interventions

Fewer studies under this engagement classification reported reasons for failure as compared to success. Among the reasons cited for failure, those related to not accounting for contextual constraints were the most consistently reported followed by implementation challenges. Most of the references to implementation challenges or failures were made by Olken 2014 and its associated studies which discussed multiple implementation barriers to a block-level grant intervention. Though less consistently reported, some studies also mentioned issues with intervention features or study design as the reasons for intervention failure. Highlighting the failure of community engagement, Olken 2014 (Grayman 2013) discussed how elite capture wherein local community leaderships and elites leveraged the intervention to retain social standing became an issue for the programme.

<table>
<thead>
<tr>
<th>Reference</th>
<th>Year</th>
<th>Type</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Okoli 2014</td>
<td>2018</td>
<td>Communication failure</td>
<td>Experiences of the beneficiaries with the programme point to problems with programme communication as they had a poor understanding of the aims of the CCT programme aside from receiving its immediate cash benefits and as such might revert to their original behaviours after the programme ends.</td>
</tr>
<tr>
<td>Baba-Ari 2018</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Study design</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Domek 2019</td>
<td>-</td>
<td>Selection bias in study enrollment</td>
<td>This may have been partly because our study population had higher baseline immunization coverage than we were adequately powered to assess. These higher than expected completion rates were likely due in part to a selection bias in enrolling children as they presented for their first immunization visit, which would have unintentionally excluded children who either presented significantly delayed or not at all for vaccines.</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Community engagement in implementation autonomy of interventions</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Supplemental material placed on this supplemental material which has been supplied by the author(s) BMJ Publishing Group Limited (BMJ) disclaims all liability and responsibility arising from any reliance on the information contained therein. BMJ Open, et al. Jain M, doi: 10.1136/bmjopen-2022-061568.
Supplementary Figure 97: Reasons for failure of interventions with engagement in implementation autonomy

Supplementary Table 16: Characteristic quotes of themes identified related to reasons for failure for interventions that used community engagement in implementation.

<table>
<thead>
<tr>
<th>Primary paper</th>
<th>Paper from which citation comes</th>
<th>Sub-theme demonstrated</th>
<th>Quote</th>
<th>Page number</th>
</tr>
</thead>
<tbody>
<tr>
<td>Oyo-Ila 2020</td>
<td>-</td>
<td>Low implementation fidelity</td>
<td>The non-impact on up-to-date vaccination could have been accounted for by the weak link in the intervention caused by not sharing data directly with community members as planned.</td>
<td>38</td>
</tr>
<tr>
<td>Olken 2014</td>
<td>World Bank 2018</td>
<td>Staff turnover</td>
<td>High village-level staff turnover may partially explain why Generasi staff in the villages are so unfamiliar with the bonus system. The entire Generasi team in Desa Lelaok (Petis subdistrict, Pamekasan district) and the PK in Desa Rampe (Nelle subdistrict, Lembata) had never heard of the bonus system, but all of them had started within the past year.</td>
<td>38</td>
</tr>
</tbody>
</table>
Webster 2019 - Competing priorities of health workers

Although the proportion of respondents reporting a home visit by a VHT decreased from baseline to endline in both control and intervention clusters the proportion of visits where immunization was the reason for the visit increased. The decrease in VHT home visits may be due in part to observations that in the latter months of the intervention, VHT services were sought by other organisations involved in the refugee response in some of the project areas, who paid VHTs higher allowances.

Not accounting for existing constraints or uncontrollable trends

Okeke 2017 - Resource constraints or scarcity

Lack of electricity and water were also frequently cited as problems.

Okeke 2017 - Poor quality of services or infrastructure

The data suggest that part of the reason why the program did not have larger impacts is that other dimensions of quality did not improve. For example, clinic infrastructure in many cases remained poor, as did availability of drugs and supplies.

Intervention features

Webster 2019 - Intervention design failure

The main mechanisms of change identified were improved accessibility to immunization services through increased numbers of vaccination outreaches, increased VHT motivation through monthly allowances together with social motivation in outdoing each other in defaulter tracing, use of community resources to achieve outcomes that is the support of local community leadership; and increased interaction between health workers and HCWs and VHTs at monthly VHT/HCW meetings. These identified mechanisms of change were predominantly due to supplementary activities that were implemented in both intervention and control clusters.

Study design
Reasons for heterogeneous impacts

A small subset of studies is also coded under the theme heterogeneous impacts. Even though they reported overall significant positive quantitative findings, these studies discussed their heterogeneous or non-uniform results with respect to not achieving the intended impact among certain subgroups, study sites or reported null results for some of the immunisation-related outcomes. Each of these studies offer a slightly different perspective on their heterogeneous results. For instance, Adamu et al. 2019 and Carnell et al. 2014 postulate that the characteristics of individuals involved in the intervention implementation, contextual factors or implementation processes could have led to partial intervention success or measurable positive gain in only select study sites. Interestingly, both Banerjee et al. 2020 and Gibson et al. 2017 highlight very different challenges with respect to their interventions comprising incentives to caregivers. Banerjee et al. describe how incentives designed to motivate caregivers did not work on Muslim minorities due to high vaccine hesitancy which stemmed from the belief that vaccination causes sterilisation. On the other hand, the Gibson study reported that their SMS reminders did not work in the absence of incentives, implying that reminders alone were insufficient to encourage behaviour change among caregivers and result in immunisation uptake. This shows that the benefit of providing incentives may not be consistent across different contexts and uptake may be influenced by other socio-economic factors. Both studies further elaborate the limits to which incentives can address deep rooted mistrust of vaccination among certain subgroups. Overall, the discussion on these studies provide important insights about how external factors, implementation variability or invalid theory of change assumptions may contribute to partial intervention success or failure. We also looked through the literature for unintended impacts but there was insufficient information to draw any broader conclusions.

Uptake and fidelity challenges

Uptake and fidelity challenges were generally identified by authors through primary data or personal experience and knowledge of the intervention. Because this theme relates to how the project was implemented, we expect that authors are a good source of this information. Individual interviews and focus group discussions supported the findings. Themes were broadly consistent in the full analysis and in sensitivity analysis in which on the qualitative papers with a quality assessment score greater than 20 were included. However, there are several instances in which a single sub-theme is dominated by one or two studies. These are indicated in Supplementary Table 17 and discussed in the text. Given this, we believe that the evidence on uptake and fidelity challenges has only moderate strength.

The most consistently reported challenges were related to administration, mobilization, fidelity, and contamination (Supplementary Figure 98; Supplementray Table 17). Administrative challenges were cited consistently across all three types of interventions, but their nature varied. Although technical issues and politics plagued interventions that used community engagement in the design of interventions or as the intervention, these issues were not common in interventions that used engagement in their implementation. Staffing was a challenge in interventions that used engagement in implementation or as the intervention. Implementation fidelity was mentioned as a challenge by all three types of interventions. In all cases, the most consistent issue was that realities on the ground forced changes to the intervention. The issue of mobilizing people into the intervention appears to be common; however, this category is dominated by two papers (Pereze 2020 and Olken 2014) which cite a variety of mobilization challenges. As such, the evidence on this theme is somewhat weak. Contamination was also consistently reported. For the most part, it was mentioned due to its absence. Interventions whose goal was
engagement discussed this topic much more than interventions with engagement in their design or implementation.

**Supplementary Figure 98: Summary figure: uptake and fidelity challenges**

Supplementary Table 17: Comparison of the challenges faced by interventions using different types of community engagement. Numbers reflect relative frequency.

<table>
<thead>
<tr>
<th>Administrative challenges</th>
<th>Community engagement as the intervention</th>
<th>Community engagement in the design of the intervention</th>
<th>Community engagement in the implementation of the intervention</th>
</tr>
</thead>
<tbody>
<tr>
<td>Technical issues</td>
<td>Technical issues</td>
<td>Funding disbursement</td>
<td></td>
</tr>
<tr>
<td>Politics</td>
<td>Politics</td>
<td>Communication across stakeholders</td>
<td></td>
</tr>
<tr>
<td>Coordination</td>
<td>Bureaucracy</td>
<td>Staffing</td>
<td></td>
</tr>
<tr>
<td>Staffing</td>
<td>Security</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mobilization challenges</td>
<td>Recruiting people into the intervention due to work, distance, and competing priorities*</td>
<td>Recruiting facilitators and participants into the intervention due to other opportunities, miscommunication of intervention expectations, declines in incentives, caregiving obligations, and more**</td>
<td></td>
</tr>
<tr>
<td>Gender based religious and cultural norms prevented participation***</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Fidelity challenges</td>
<td>Failure to account for on the ground realities resulted in changes or failures in intervention</td>
<td>Failure to account for on the ground realities resulted in changes or failures in intervention implementation</td>
<td>Failure to account for on the ground realities resulted in changes or failures in intervention implementation</td>
</tr>
</tbody>
</table>

* Coding is presented in a present/absent format because, in some cases, authors explicitly stated that a challenge was not encountered.
Community engagement as the intervention

Discussions of administrative challenges consistently related to coordination, technical issues, staffing, and politics. The single study that reported on the absence of certain administrative issues also reported on the presence of many others. This is likely because it was a 3ie report, which tend to be quite detailed and include significant qualitative components.

Issues with the faithful implementation of these interventions were consistently the result of improper supervision and management or a failure to account for practicalities on the ground. We found one case where the administrative challenges directly led to low implementation fidelity. However, a non-trivial proportion of studies reported statistics indicating high fidelity.

The studies that discussed contamination largely did so in order to demonstrate that contamination was not a concern. Authors consistently focused on the issue of contamination between intervention and control sites. This was often addressed with vague statements. Distance was largely used to avoid contamination.

Discussion of challenges in mobilization interventions was dominated by a single source (Olken et al 2014 also discussed the engagement in implementation section). However, Rahman 2016 also discussed several mobilization issues that resulted from religious and cultural norms, largely around gender, preventing participation.
**Supplementary Figure 99: Engagement as intervention**

Supplementary Table 18: Characteristic quotes of themes identified related to uptake and fidelity challenges for interventions were community engagement.

<table>
<thead>
<tr>
<th>Primary paper</th>
<th>Paper from which citation comes</th>
<th>Sub-theme demonstrated</th>
<th>Quote</th>
<th>Page number</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>Administrative challenges</em></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Assengai et al 2018</td>
<td>Khuzwayo et al 2017</td>
<td>Coordination challenges</td>
<td>Respondents report a lack of communication about visit scheduling.</td>
<td>3</td>
</tr>
<tr>
<td>Demilew et al 2019</td>
<td>Demilew et al 2020</td>
<td>Technical and political challenges</td>
<td>This is compounded by dual reporting lines in many parts of the country, where CHWs remain linked to and receive stipends through NGO intermediaries, while being accountable to PHC facility managers.</td>
<td>23</td>
</tr>
</tbody>
</table>

The Mobile Intervention was left from being implemented because of Internet service inaccessibility that resulted from the state of emergency declared by the Ethiopian government.
<table>
<thead>
<tr>
<th>Political challenges</th>
<th>This study has faced various degrees of support from our primary government stakeholder since its inception, and this is the main reason why the timeline has been extended multiple times over the past seven years (when discussions with the Haryana government first began). We faced issues in continuity and buy-in especially when higher-ranking officers were transferred.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Absence of administrative barriers</td>
<td>Given that J-PAL South Asia was also responsible for implementation, we recruited a dedicated team with relevant experience to do this. As a result, we were acutely aware of what was happening in the field, and able to identify and resolve issues quickly.</td>
</tr>
</tbody>
</table>

**Implementation fidelity**

<table>
<thead>
<tr>
<th>Biemba et al 2016</th>
<th>Phiri et al 2017</th>
<th>Inadequate supervision</th>
<th>Supervision and mentorship is crucial to optimizing the skills of CHAs, but this study outlines that supervision is not always implemented as intended and that in the absence of regular supervision.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Morris et al 2004</td>
<td>Morris et al 2004</td>
<td>Expectations did not match realities on the ground</td>
<td>The service-level package was not implemented in accordance with protocol because no legal means could be identified of transferring resources.</td>
</tr>
<tr>
<td>Demilew et al 2019</td>
<td>Demilew et al 2020</td>
<td>Administrative challenges caused implementation challenges</td>
<td>This intervention was designed on the expectation that the text message portion could be rolled out in tandem with the ECIIN, an immunisation tracking program that planned to have HEWs report immunisation dates, mother’s data, child’s data, using text messages. Due to data issues and conflict in the region ECIIN was discontinued.</td>
</tr>
<tr>
<td>Gurley et al 2020</td>
<td>Gurley et al 2020</td>
<td>High fidelity</td>
<td>The types of screenings were conducted as planned, with 446 mothers group screenings, 445 VHND screenings, 223 men’s screenings, and 444 HTR screenings (Table 6). Within a single village, this translated to an average of 12 sessions held in mothers groups, 12 held in VHNDs, 12 held in HTR areas, and 6 held in men’s groups.</td>
</tr>
</tbody>
</table>

**Contamination**

| Moore et al 2017 | Moore et al 2017 | Absence of contamination | Our intervention and control groups were generally similar, with high coverage and fidelity to planned activities and negligible contamination. |

221
To mitigate potential contamination between intervention and control villages, we attempted to ensure that the intervention and control villages are sufficiently far apart from one another.

*Challenges related to mobilisation are not presented because this information largely came from a single paper making the identification of characteristic issues across papers impossible.

**Community engagement in the design of the intervention**

Among interventions that used community engagement in their design, the most consistently reported administrative challenge related to technical issues, largely surrounding phones and internet connections (Supplementary Table 17). However, other administrative challenges were related to combinations of politics, bureaucracy, and security concerns. The single paper that indicated a lack of administrative challenges reported successful syncing to the server, but also reported on other technological issues.

Mobilization challenges largely related to recruiting people into the intervention. Most of this information came from a 2020 (Pereze et al 2020) paper which presented information on the implementation fidelity and acceptability of an intervention that combined in-person and digital education to promote vaccination (Johri 2018). Authors describe challenges with participants being unable to attend due to work, distance, and competing priorities. Limited access to phones made mobilization difficult as well.

Acknowledgement of low implementation fidelity was consistently accompanied by discussions of why changes were made or the implications of these changes for results. Generally, changes in the intervention plan were forced due to practical constraints that had not been accounted for and were expected to result in an attenuation of impacts. Reports of high fidelity were succinct and responded to concerns a reader may be expected to have.

Contamination can occur between the intervention and control or due to ongoing activities of other programs. The former of these was reported more consistently than the latter. Authors were more likely to discuss the absence of this issue than to report upon its presence. Some offered explanations as to why contamination was not a concern or steps taken to avoid contamination. Others reported on results related to measures of contamination.
Supplementary Figure 100: Engagement in design

Supplementary Table 19: Characteristic quotes of themes identified related to uptake and fidelity challenges for interventions that used community engagement in their design

<table>
<thead>
<tr>
<th>Primary paper</th>
<th>Paper from which citation comes</th>
<th>Sub-theme demonstrated</th>
<th>Quote</th>
<th>Page number</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Administrative challenges</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Nagar et al 2020</td>
<td>Nagar et al 2020</td>
<td>Technological problems</td>
<td>The most prevalent issues reported by ANMs using the KB App included having to reenter data after an app crash (60.9%), having issues with scanning the pendant (65.2%), and issues with the application crashing (56.5%), and time required to enter data (33.3%).</td>
<td>138</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Technological problems</td>
<td>Limited phone memory occasionally created technology-related issues, with the increasing requirement to store data on the mobile phone.</td>
<td>10</td>
</tr>
<tr>
<td>Modi et al 2019</td>
<td>Modi et al 2015</td>
<td>Absence of technological problems</td>
<td>The failure rate, defined as the proportion of forms that failed to get synced with the server, was &lt;1%.</td>
<td>266</td>
</tr>
</tbody>
</table>

**Implementation fidelity**
Gruley et al 2020  Reason changes were implemented  Last, due to resource constraints, the project was not able to implement the intervention component for sharing videos from phone to phone using Bluetooth technology.

Engineer et al 2019 Implications of low fidelity for the research  Problems with implementation likely dampened any potential effect. The scheme was rolled out in phases, but there were some delays, particularly with the initial payments.

Banerjee et al 2010 High implementation fidelity  Review of records showed that of 1336 planned camps, 95% (1269) took place.

**Contamination**

Borkum et al 2014 Absence of contamination with other interventions  Because FLWs and households in the treatment and control groups received non-TBGI Ananya program interventions that were being implemented simultaneously across Begusarai, the RCT was designed to measure the value-added of TBGI beyond these other Ananya interventions. It was not designed to measure the impact of TBGI introduced in isolation.

Nagar et al 2018 Steps to avoid contamination  a cluster randomised approach further allowed non-contamination of social signaling interventions within the contacts of a given village

Johri et al 2018 Pereze et al 2020 Little contamination observed  According to the analysis of the records, only people who belonged to the intervention group attended the community meetings. The people who communicated with or received calls from the IVR platform were from the intervention group, except for one (1/166) control group member.

*Challenges related to mobilisation are not presented because this information largely came from a single paper making the identification of characteristic issues across papers impossible.

**Community engagement in the implementation autonomy of the intervention**

*Administrative challenges* in interventions that used engagement in their implementation largely related to funding disbursement, communication across stakeholders, staffing issues (Supplementary Table 17). One study (Bjorkman et al 2009) reported that they did not believe administrative differences affected results. This was
coded as the absence of an administrative challenge, but does not mean that administrative issues were not encountered.

*Implementation fidelity* appears to have been more of a challenge for these interventions than for those that leveraged community engagement in their design. These issues were consistently related to the intervention not functioning as expected. In some cases, this was because the reality on the ground did not reflect expectations. However, two studies that reference challenges with implementation fidelity also report upon successes in implementing according to plan.

All reports of challenges *mobilising* participants into interventions that were implemented by the community came from a paper related to an intervention that used block grants to support maternal and child health and education (Olken 2014). Challenges were encountered in the recruitment of both facilitators and participants. These related to other opportunities, miscommunication of intervention expectations, declines in incentives, caregiving obligations, and more. However, Herrara-Almanza 2018 report that community health workers felt strong responsibility for the wellbeing of their communities and attrition was low.

Interventions that used engagement in their implementation also experienced significant *budgetary constraints*. Almost universally, these constraints were the result of inadequate or delayed pay. Often, this resulted in activities not being implemented.

**Supplementary Figure 101: Engagement in implementation autonomy of the intervention**

**Supplementary Table 20:** Characteristic quotes of themes identified related to uptake and fidelity challenges for interventions that used community engagement in implementing the intervention.

<table>
<thead>
<tr>
<th>Primary paper</th>
<th>Paper which comes from citation</th>
<th>Sub-theme demonstrated</th>
<th>Quote</th>
<th>Page number</th>
</tr>
</thead>
<tbody>
<tr>
<td>Olken 2014</td>
<td>Febriany et al</td>
<td>Delays in funding</td>
<td>The most frequent complaint by FDs</td>
<td>14</td>
</tr>
<tr>
<td>2011</td>
<td>Mayumana et al 2017</td>
<td>Chimbutu et al 2011</td>
<td>Stakeholder communication problem</td>
<td>The NTPI’s programme, however, did not start in 2007 as planned… The major reason for this is on the implementation challenges between the two partners. The MOHSW preferred an accelerated approach while NORAD preferred an implementation framework to be designed by the Ifakara Health Institute (IHI).</td>
</tr>
<tr>
<td>Oyo-Ita 2020</td>
<td>Oyo-Ita 2020</td>
<td>Staffing and bureaucracy challenges</td>
<td>Low staffing of facilities constrained health staff from deploying the defaulters’ register. Some complained that they had many registers other than the defaulters’ register to fill.</td>
<td></td>
</tr>
</tbody>
</table>

**Implementation fidelity**

| Oyo-Ita 2020 | Oyo-Ita 2020 | Inconsistent implementation | The community engagement did not use town hall meetings as planned. Only one town hall meeting was held. |
| Okeke et al 2017 | Okeke et al 2017 | Expectations did not match realities on the ground | Contrary to the role of the states outlined in the MOU, policymakers in Enugu state reported playing no role in the implementation of the scheme. One suggested that the division of responsibility drawn up in the MOU did not align with the setup of health care in Nigeria given that the state government has responsibility for secondary care only. |
| Mayumana et al 2017 | Olafsdottir et al 2014 | Intervention did not function as expected | Firstly, the work load on staff is considerably higher than they are contracted for, which is not reflected in the remuneration; payments are delayed and overtime and eligible allowances are not always paid causing demotivation among staff. |
| | | Intervention was implemented according to clear standards | The direct and transparent payment of funds, as well as more frequent contact with their managers was reported to enhance trust and improve the relationship between health workers and... |
their managers.

**Budgetary constraints**

Okeke et al. 2017

Pay was severely delayed

Some local government don’t have the money to pay and they have not paid for a period of time. Most of them [midwives] have even finished their service and gone out without receiving any penny from the local government.

Mayumana et al. 2017

Activities were suspended

According to the qualitative data one of the obstacles for providing more frequent supervision visits was a lack of financial resources:

We do supervision every quarter. Our plan is to do supervision monthly but due to limited budget we haven’t yet done this.

*Challenges related to mobilization are not presented because this information came from a single paper making the identification of characteristic issues across papers impossible.

**Sensitivity analysis for qualitative evidence**

We carried out a sensitivity analysis in which we considered only the 17 qualitative studies associated with 12 impact evaluations that had a risk of bias assessment score of 20 or higher. Broadly, the themes that emerged from this analysis were consistent with those of the full analysis. We did not have enough qualitative evidence to conduct a sensitivity analysis by engagement type. An important caveat of this analysis is that nearly half of the papers rated as high “quality” were associated with three impact evaluations. The panel figures below provide comparison of the emerging themes in the full analysis versus the sensitivity analysis by comparing hierarchy charts related to barriers, facilitators, reasons for project success or failure, and uptake and fidelity challenges.

**Barriers to immunisation**

The most common barriers were broadly similar across the full analysis and the sensitivity analysis. *Behavioural, social and practical barriers faced by caregivers* were the most consistently reported barriers followed by *constraints in delivery of immunisation services* (Supplementary Figure 102). The most common barriers faced by caregivers were unavailability of immunisation services, financial constraints and social norms.

![Supplementary Figure 102 (a): Full analysis](image1)

![Supplementary Figure 102 (b): Sensitivity analysis](image2)
Facilitators of immunisation

The most common facilitators of immunisation were broadly similar across the full analysis and the sensitivity analysis. Behavioural, social and practical factors faced by caregivers were the most consistently reported facilitators of immunisation in a given context followed by good delivery of immunisation services (Supplementary Figure 103). The most consistently reported facilitators were availability of immunisation services, enabling social norms and caregivers’ understanding of importance of immunisation.

Supplementary Figure 103 (a): Full analysis  (b): Sensitivity analysis

Reasons for intervention success

The most consistently reported reasons for success were broadly similar across the full analysis and the sensitivity analysis. Success was consistently attributed to intervention features including leadership and supportive supervision, community engagement, health worker training, incentives and customisation to local context. Though not as common, success was also attributed to existing or changing favourable characteristics within a given context such as positive participant views, enabling social norms and availability of and access to good quality health services. Given that a very small subset of papers were included in the sensitivity analysis, a few themes that emerged in the full analysis were missing in the sensitivity analysis. Supplementary Figure 104 provides the hierarchy charts comparing the two analyses.

Supplementary Figure 104 (a): Full analysis  (b): Sensitivity analysis

Reasons for intervention failure

The most consistently reported reasons for failure were broadly similar across the full analysis and the sensitivity analysis. The most consistent reasons for failure were attributed to accounting for contextual constraints and inadequate intervention features. Social norms, health system related issues, political or civil unrest and high baseline coverage were some of the most common contextual reasons for failure. Among the intervention features, inadequate duration, frequency or exposure to the intervention were the most notable reasons for failure. Though not as prominently reported as compared to the full analysis, implementation challenges were also noted to have caused intervention failure in the sensitivity analysis. Among these,
disruption due to inadequate implementation instructions and competing priorities of health workers were noted as some of the prevalent challenges.

**Uptake and fidelity challenges**

The most consistently reported uptake and fidelity issues were broadly similar across the full analysis and the sensitivity analysis. The most common challenges consistently reported challenges were related to administration, mobilization, fidelity, contamination and budget constraints. Administrative challenges were cited consistently. Similar to the full analysis, the primary most consistent issue was that realities on the ground forced changes to the intervention.
Appendix 15: Synthesis of cost evidence

This section reports on the results of the analysis of the inventory of included studies (see Table 21) and of the estimates of the non-vaccine cost per dose of interventions to increase absolute immunisation coverage by one percent (see Table 22).

Characteristics of included studies

Twenty-two of the evaluations selected for inclusion in the systematic review reported some type of cost analysis. Of these, 18 used an experimental design and the remaining 4 used quasi-experimental methods to identify the impact of treatment.

Of the 22 studies: 20 studies reported a total cost of the intervention; 12 included a cost-effectiveness analysis; three included cost-efficiency analysis; two studies included a cost-benefit analysis and two analysed cost per quality-adjusted life year (QALY).

Table 21: Characteristics of studies including any type of cost analysis.

<table>
<thead>
<tr>
<th>Author, Date</th>
<th>Country</th>
<th>Evaluation design</th>
<th>Intervention exposure (months)</th>
<th>Total Cost</th>
<th>Cost-efficiency Analysis</th>
<th>Cost-effectiveness Analysis</th>
<th>Cost-benefit Analysis</th>
<th>Cost per QALY or DALY</th>
</tr>
</thead>
<tbody>
<tr>
<td>Andersson, 2013</td>
<td>Pakistan</td>
<td>Experimental</td>
<td>8</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Banerjee, 2010</td>
<td>India</td>
<td>Experimental</td>
<td>18</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Banerjee, 2020</td>
<td>India</td>
<td>Experimental</td>
<td>14</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Björkman, 2009</td>
<td>Uganda</td>
<td>Experimental</td>
<td>0.17</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Borkum, 2020</td>
<td>India</td>
<td>Experimental</td>
<td>12</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Demilew, 2020</td>
<td>Ethiopia</td>
<td>Experimental</td>
<td>17</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Gurley, 2020</td>
<td>India</td>
<td>Experimental</td>
<td>11</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Johri, 2009</td>
<td>India</td>
<td>Experimental</td>
<td>3</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Manoj, 2020</td>
<td>India</td>
<td>Experimental</td>
<td>4</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Modi, 2019</td>
<td>India</td>
<td>Experimental</td>
<td>12</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>More, 2017</td>
<td>India</td>
<td>Experimental</td>
<td>24</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Morris, 2020</td>
<td>Honduras</td>
<td>Experimental</td>
<td>24</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Nagar, 2020</td>
<td>India</td>
<td>Experimental</td>
<td>20</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Olken, 2014</td>
<td>Indonesia</td>
<td>Experimental</td>
<td>24</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Oyo-Ito, 2020</td>
<td>Nigeria</td>
<td>Experimental</td>
<td>18</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Pramanik, 2018</td>
<td>India</td>
<td>Experimental</td>
<td>13</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Seth, 2010</td>
<td>India</td>
<td>Experimental</td>
<td>9.7</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Webcer, 2019</td>
<td>Uganda</td>
<td>Experimental</td>
<td>12</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Admassie, 2017</td>
<td>Ethiopia</td>
<td>Quasi-experimental</td>
<td>48</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Carnell, 2014</td>
<td>Ethiopia</td>
<td>Quasi-experimental</td>
<td>48</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Findley, 2004</td>
<td>Nigeria</td>
<td>Quasi-experimental</td>
<td>24</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Saggurti, 2009</td>
<td>India</td>
<td>Quasi-experimental</td>
<td>2</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
</tbody>
</table>

By country. Fourteen of the community participation interventions to improve child immunisation studies were conducted in Asian countries: of these, 12 studies took place in India; and one each in Pakistan and Indonesia. One study was conducted in a Central American country, Honduras and seven studies were conducted in African countries: three in Ethiopia; and two each in Nigeria and Uganda
Supplementary Figure 107: Number of evaluations with any type of cost analysis
The production of cost evidence in published evaluations of community participation interventions to improve child immunisation in low- and middle-income countries was uneven over the years. The first recorded evaluation with any cost analysis was published in 2004. Between 2009 and 2019, no more than three of the identified evaluations included any type of cost analysis in any given year. In 2020, the largest number of evaluations, eight, were published in a single year. The eight evaluations published in 2020 were required to generate cost analysis in partial fulfillment of the grant requirements in 3ie’s Innovations in Increasing Immunisation Evidence Programme. Funding for these studies was provided by the Bill and Melinda Gates Foundation.

Non-vaccine cost per dose of intervention to increase immunisation coverage by one percent
Of the 22 included studies reporting cost and impact information, we estimate the non-vaccine cost per dose of intervention to increase immunisation coverage by one percent for 14. Because several studies report results by treatment arm, we report results for a total of 17 treatment arms. These results are summarised in detail in Table 22.

Three studies and one treatment arm were dropped from the analysis. In the case of Johri, 2020 this was because the proportion of children vaccinated was not reported as an outcome of this evaluation. Morris 2020 was dropped because it is unclear how to combine the PRAF II costs reported in a separate paper with intervention as described in Morris, 2020, in addition, the evaluation found a null result on the immunisation outcomes. Findley, 2013 was dropped because the reported cost information is insufficient for generating a total cost estimate or cost per vaccine dose, and “immunisation status” is not clearly defined. In the case of Banerjee, 2020, no coverage change information was reported for the Gossip seeds treatment arm which was dropped from the analysis as a result.

Six studies, Findley, 2004; Pramanik, 2018; Webster, 2019; Oyo-Ita, 2020, Modi, 2019; and Demilew, 2020 reported negative or null results and were excluded from the analysis of non-vaccine cost per dose of intervention to increase immunisation coverage by one percent.

The median non-vaccine cost per dose of intervention to increase immunisation coverage by one percent for the 14 estimates was US $3.68 (all costs are reported in 2019 US dollars). In comparison, the average was US $46.60 which is driven by three observations, Olken 2014 (US $23.10); Carnell, 2014 (US $641.08); and Gurley, 2020 (US $29.98).
Supplementary Figure 108: Non-vaccine cost per dose of intervention to increase immunization coverage by one percent

Supplementary Figure 108 illustrates that the range of estimates varied from a minimum of US $0.88 to a maximum of US $29.98. The lowest cost per vaccine dose was observed in Banerjee, 2010 for the “camps with incentives” intervention at US $0.88 per vaccine dose to increase immunisation coverage by one percent. The highest cost was observed in Carnell, 2014 at $641.08 - where the five-year intervention included system-building activities such as building and staffing a health post. In addition, the transparency of cost reporting in Carnell is very low, which further contributes to uncertainty in this cost estimate. Indeed, Munk et. Al (2019) dropped this study from their systematic review of costs because they judged the available cost information to be insufficient to calculate the cost per vaccine dose. Since Carnell is an outlier observation, we take the intervention with the second-highest cost per incremental outcome which was reported in Gurley, 2020 at US $29.98.
Supplementary Table 22: Detailed estimates of non-vaccine intervention cost per dose and absolute coverage change

<table>
<thead>
<tr>
<th>Author Date</th>
<th>Source</th>
<th>Total Cost USD</th>
<th>Treatment</th>
<th>Outcome</th>
<th>Number of doses</th>
<th>2019 USD</th>
<th>Nature-Randomized</th>
<th>Non-vaccine intervention cost per vaccine dose</th>
<th>Absolute change in coverage due to intervention</th>
<th>Non-vaccine cost per dose of intervention to increase intervention coverage</th>
<th>Cost rate</th>
<th>Immunization rates</th>
<th>Best</th>
</tr>
</thead>
<tbody>
<tr>
<td>Jain et al. 2022</td>
<td>Vaccine</td>
<td>$15.156</td>
<td>Vaccination</td>
<td>Full immunization</td>
<td>5 per visit</td>
<td>$4.39</td>
<td>95%</td>
<td>Reduced proportion of unimmunized children</td>
<td>44% increase in coverage</td>
<td>8% increase in coverage, 4% increase in coverage</td>
<td>$0.69</td>
<td>90%</td>
<td></td>
</tr>
<tr>
<td>Das et al. 2020</td>
<td>Reduce</td>
<td>$224.569</td>
<td>Vaccination</td>
<td>Full immunization</td>
<td>6 per visit</td>
<td>$23</td>
<td>95%</td>
<td>Reduced proportion of unimmunized children</td>
<td>44% increase in coverage</td>
<td>8% increase in coverage, 4% increase in coverage</td>
<td>$0.69</td>
<td>90%</td>
<td></td>
</tr>
<tr>
<td>Sen et al. 2018</td>
<td>India</td>
<td>$2,738</td>
<td>Vaccination</td>
<td>Full immunization</td>
<td>4 per visit</td>
<td>$4</td>
<td>95%</td>
<td>Reduced proportion of unimmunized children</td>
<td>44% increase in coverage</td>
<td>8% increase in coverage, 4% increase in coverage</td>
<td>$0.69</td>
<td>90%</td>
<td></td>
</tr>
<tr>
<td>Benavides et al. 2018</td>
<td>India</td>
<td>$24,940</td>
<td>Vaccination</td>
<td>Full immunization</td>
<td>3.5 per visit</td>
<td>$151.44</td>
<td>95%</td>
<td>Reduced proportion of unimmunized children</td>
<td>44% increase in coverage</td>
<td>8% increase in coverage, 4% increase in coverage</td>
<td>$0.69</td>
<td>90%</td>
<td></td>
</tr>
<tr>
<td>Benavides et al. 2018</td>
<td>India</td>
<td>$15,403</td>
<td>Vaccination</td>
<td>Full immunization</td>
<td>2 per visit</td>
<td>$9.2</td>
<td>95%</td>
<td>Reduced proportion of unimmunized children</td>
<td>44% increase in coverage</td>
<td>8% increase in coverage, 4% increase in coverage</td>
<td>$0.69</td>
<td>90%</td>
<td></td>
</tr>
<tr>
<td>Pande et al. 2018</td>
<td>India</td>
<td>$78,821.4</td>
<td>Vaccination</td>
<td>Full immunization</td>
<td>1 per visit</td>
<td>$31.4</td>
<td>95%</td>
<td>Reduced proportion of unimmunized children</td>
<td>44% increase in coverage</td>
<td>8% increase in coverage, 4% increase in coverage</td>
<td>$0.69</td>
<td>90%</td>
<td></td>
</tr>
<tr>
<td>Jain et al. 2018</td>
<td>India</td>
<td>$50,466.6</td>
<td>Vaccination</td>
<td>Full immunization</td>
<td>0.5 per visit</td>
<td>$13.3</td>
<td>95%</td>
<td>Reduced proportion of unimmunized children</td>
<td>44% increase in coverage</td>
<td>8% increase in coverage, 4% increase in coverage</td>
<td>$0.69</td>
<td>90%</td>
<td></td>
</tr>
<tr>
<td>Author(s)</td>
<td>Count</td>
<td>Title/Conc ID</td>
<td>Treatment</td>
<td>Pre/Post</td>
<td>Number of subjects</td>
<td>Baseline age (years)</td>
<td>Baseline value</td>
<td>Baseline &amp; intervention</td>
<td>Change (non-significant)</td>
<td>Change (significant)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>----------</td>
<td>-------</td>
<td>--------------</td>
<td>-----------</td>
<td>---------</td>
<td>-------------------</td>
<td>---------------------</td>
<td>--------------</td>
<td>------------------------</td>
<td>----------------------</td>
<td>---------------------</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Jain M. et al.</td>
<td>BMJ Open</td>
<td>1202526.56</td>
<td></td>
<td></td>
<td>9</td>
<td>30-69</td>
<td>50</td>
<td>40</td>
<td>10</td>
<td>20</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>BMJ Open, et al.</td>
<td>Jain M.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Country</td>
<td>Year</td>
<td>Authors</td>
<td>Intervention</td>
<td>Cost</td>
<td>Cost Type</td>
<td>Outcome</td>
<td>Notes</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>---------</td>
<td>------</td>
<td>---------</td>
<td>--------------</td>
<td>------</td>
<td>-----------</td>
<td>---------</td>
<td>-------</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Uganda</td>
<td>2019</td>
<td>Bijur and Banerjee</td>
<td>Back of the envelope (protocol)</td>
<td>$182,003.09</td>
<td>$13.81</td>
<td>0%</td>
<td>Baseline immunization not reported 20.8% - 50.05% is the range of baseline means, depending on the district for a child fully vaccinated 20.8% - 50.05% is the range of baseline means, depending on the district for a child fully vaccinated</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>India</td>
<td>2020</td>
<td>Webster</td>
<td>High slope incentives</td>
<td>$242,595.94</td>
<td>$21.41</td>
<td>0%</td>
<td>Adjusted percent increase in fully immunised 1,055.51:1 intervention control group, excludes healthcare system-building costs</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>India</td>
<td>2020</td>
<td>Nagar</td>
<td>Low slope incentives</td>
<td>$367,078.25</td>
<td>$21.14</td>
<td>0%</td>
<td>Adjusted percent change in fully immunised 1,055.51:1 intervention control group, excludes healthcare system-building costs</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>India</td>
<td>2020</td>
<td>Gurley</td>
<td>Gossip seeds</td>
<td>$76,073.73</td>
<td>0%</td>
<td>0%</td>
<td>Unadjusted proportion, endline (intervention - control). Difference is significant at p &lt; 0.05.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>India</td>
<td>2020</td>
<td>Banerjee</td>
<td>Tracking system: The Khushi Baby platform, a wearable digital health record (pendant) and with patient-specific, dialect-specific voice reminders</td>
<td>$275,702.59</td>
<td>$11.68</td>
<td>0%</td>
<td>Incremental to no intervention control group, excludes healthcare system-building costs</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>India</td>
<td>2020</td>
<td>Crook</td>
<td>PATH - community-led videos</td>
<td>$175,603.81</td>
<td>$29.98</td>
<td>0%</td>
<td>Incremental to no intervention control group, excludes healthcare system-building costs</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ethiopia</td>
<td>2014</td>
<td>Carrell</td>
<td>Three pillars approach (the health project: strengthening health systems, improving health workers’ performance, and engaging the community)</td>
<td>$273,165,777.5</td>
<td>$641.08</td>
<td>0%</td>
<td>Incremental to no intervention control group, excludes healthcare system-building costs</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>India</td>
<td>2019</td>
<td>Modi</td>
<td>A mobile-phone app to assist ASHAs and health providers to schedule and track home visits and follow-up care</td>
<td>$160,081.37</td>
<td>0%</td>
<td>0%</td>
<td>For a child fully vaccinated 20.8% - 50.05% is the range of baseline means, depending on the district for a child fully vaccinated</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Uganda</td>
<td>2019</td>
<td>Webster</td>
<td>Improve dosa platform and the Fifth Child, a community engagement strategy to increase vaccination coverage</td>
<td>$224,224.18</td>
<td>0%</td>
<td>0%</td>
<td>For a child fully vaccinated 20.8% - 50.05% is the range of baseline means, depending on the district for a child fully vaccinated</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>