Jade Benjamin-Chung
Jaynal Abedin
David Berger
Ashley Clark
Lauren Falcao
Veronica Jimenez
Eugene Konagaya
Diana Tran
Benjamin F Arnold
et al.

The identification and measurement of health-related spillovers in impact evaluations

A systematic review

October 2015

Systematic Review 22

Health

International Initiative for Impact Evaluation

About 3ie

The International Initiative for Impact Evaluation (3ie) is an international grant-making NGO promoting evidence-informed development policies and programmes. We are the global leader in funding, producing and synthesising high-quality evidence of what works, for whom, why and at what cost. We believe that better and policy-relevant evidence will make development more effective and improve people's lives.

3ie systematic reviews

3ie systematic reviews appraise and synthesise the available high-quality evidence on the effectiveness of social and economic development interventions in low- and middle-income countries. These reviews follow scientifically recognised review methods, and are peer-reviewed and quality assured according to internationally accepted standards. 3ie is providing leadership in demonstrating rigorous and innovative review methodologies, such as using theory-based approaches suited to inform policy and programming in the dynamic contexts and challenges of low- and middle-income countries.

About this review

The identification and measurement of health-related spillovers in impact evaluations: a systematic review, was submitted in partial fulfilment of the requirements of grant SR4.1084 issued under Systematic Review Window 4. This review is available on the 3ie website. 3ie is publishing this report as received from the authors; it has been formatted to 3ie style. 3ie has published a summary report of this review, designed for use by decision-makers, which is available here.

All content is the sole responsibility of the authors and does not represent the opinions of 3ie, its donors or its board of commissioners. Any errors are also the sole responsibility of the authors. Comments or queries should be directed to the corresponding author, Jade Benjamin-Chung, jadebc@berkeley.edu.

Major funding for this systematic review was provided by UK aid, the Bill & Melinda Gates Foundation and the Hewlett Foundation. A complete listing of all of 3ie's donors can be found on the 3ie website.

Suggested citation: Benjamin-Chung, J, Abedin, J, Berger, D, Clark, A, Falcao, L, Jimenez, V, Konagaya, E, Tran, D, Arnold, B, Hubbard, A, Luby, S, Miguel, E and Colford, J, 2015. *The identification and measurement of health-related spillovers in impact evaluations: a systematic review*, *3ie Systematic Review 22*. London: International Initiative for Impact Evaluation (3ie).

3ie systematic review executive editors: Philip Davies and Beryl Leach
Managing editor: Deepthy Menon
Technical editor: Hugh Waddington
Production manager: Pradeep Singh
Cover design: John F McGill and Akarsh Gupta

© International Initiative for Impact Evaluation (3ie), 2015, 2016

The identification and measurement of health-related spillovers in impact evaluations: a systematic review

Jade Benjamin-Chung
University of California, Berkeley

Jaynal Abedin
icddr,b

David Berger
University of California, Berkeley

Ashley Clark
University of California, Berkeley

Lauren Falcao
University of California, Berkeley

Veronica Jimenez
University of California, Berkeley

Eugene Konagaya
University of California, Berkeley

Diana Tran
University of California, Berkeley

Benjamin F Arnold
University of California, Berkeley

Alan Hubbard
University of California, Berkeley

Stephen P Luby
Stanford University

Edward Miguel
University of California, Berkeley

John M Colford, Jr
University of California, Berkeley

3ie Systematic Review 22

October 2015

International Initiative for Impact Evaluation

Summary

Background: Many interventions delivered to improve global health may benefit not only direct recipients but also people in close physical or social proximity to them. These “spillover effects” are of increasing interest across disciplines. However, methods to estimate spillovers have not been developed systematically or standardized across academic disciplines.

Objectives: To summarize published methods used to estimate health-related spillover effects, summarize existing literature on these effects, and provide recommendations for the design, analysis, and reporting of future studies in which spillover effects are to be measured.

Search methods: We searched 19 electronic databases for articles published before 2014 and hand-searched titles from 2010-2013 in five journals in relevant fields.

Selection criteria: 1) Studies in low- or middle-income countries, 2) quantitative studies evaluating an intervention, 3) studies measuring health outcomes, 4) studies clearly articulating a counterfactual for the spillover parameter that was estimated.

Data collection and analysis: At least one team member reviewed each record retrieved for relevance. We classified the spillover parameters estimated in each included study and compared results within spillover parameter classes and intervention type. We adapted the CONSORT checklist for reporting randomized trials to include estimation of spillovers in randomized or observational studies.

Main results: We reviewed 34,042 titles, 12,836 abstracts, and 775 full texts published until 2014. Fifty-four studies conducted in 21 low- and middle-income countries met our inclusion criteria. Studies evaluated a wide range of interventions including vaccines (n=22 studies), mass drug administration for infectious disease control (n=7), cash transfers (n=5), and women’s education and empowerment programs (n=2). We identified 22 different spillover parameters estimated in the included studies. In general, the proportion of individuals receiving an intervention in a population was associated with improved health outcomes. In our assessment of the overall quality of study evidence, 6 of the 54 included studies (11%) had high quality evidence, 30 (56%) had moderate quality, 12 (22%) had low quality, and six (11%) had very low quality evidence. We found evidence of publication bias for certain spillover estimates but not for total or direct effects.

Conclusions: A wide range of methods was used across academic disciplines to estimate health spillovers in low- and middle-income countries. The strongest evidence for spillovers was present for vaccines and mass drug administration to control parasites. We recommend that future studies pre-specify spillover measurement and report spillovers using our checklist to ensure thorough reporting and allow for greater comparability of spillover findings.

Funding: Funding for this systematic review was provided by the International Initiative for Impact Evaluation (3ie).

Contents

Summary ..... **i**
List of tables and figures……..**iii**
1. Background **1**
2. Objectives **3**
3. Methodological synthesis **4**
3.1. Theories of spillover mechanisms 4
3.2. Methods for estimating spillovers 5
4. Methods **19**
4.1. Protocol and registration 19
4.2. Criteria for considering studies for this review 19
4.3. Search Methods for Identification of Studies 21
4.4. Data Collection and Management 21
4.5. Critical appraisal and Risk of Bias assessment 22
4.6. Unit of Analysis Issues 23
4.7. Assessment of study dependence 23
4.8. Synthesis of findings 23
4.9. Additional analyses 24
5. Results **25**
5.1. Description of Studies 25
5.2. Methodological Quality and Risk of Bias in Included Studies 27
5.3. Synthesis of Results 29
5.4. Publication Bias 45
5.5. Additional Analyses 46
6. Discussion **47**
6.1. Summary of Main Results 47
6.2. Overall Completeness and Applicability of Evidence 49
6.3. Quality of the Evidence 49
6.4. Limitations and Potential Biases in the Review Process 50
7. Conclusions **51**
7.1. Implications for Practice and Policy 51
7.2. Implications for Research 51
7.3. Recommendations 53
References……..**57**
8. Plans for Updating the Review **68**
9. Acknowledgments **69**
10. The Review Team and Contributions of Authors **70**
11. Statement Concerning Conflict of Interest **71**
12. Tables and Figures **72**
Appendix 1: Search strategy **105**
Appendix 2. Electronic databases searched **131**
Appendix 3: Risk of bias assessment **133**
Appendix 4. Rationale for the classification of each type of study design for estimating spillover effects as low, moderate, or high quality. **136**
Appendix 5. Comprehensive validity assessment **139**
Appendix 6. Reasons for exclusion of full texts **140**
Appendix 7. Overall quality of evidence for each included study **141**

List of tables and figures

Table 1: Primary design, scale, and mechanism for spillover parameter classes 72
Table 2a: Modifications to Cochrane “GRADE” tool to incorporate spillover assessment: classification of studies’ underlying methodology..... **Error! Bookmark not defined.**
Table 2b: Modifications to Cochrane “GRADE” tool to incorporate spillover assessment: factors that may increase or decrease the quality level of a body of evidence**Error! Bookmark not defined.**Table 3: Characteristics of studies meeting inclusion criteria 75
Table 4: Study designs by academic discipline 81
Table 5: Intervention types by academic discipline.....82
Table 6: Intervention types by reason for cluster vs. individual level treatment allocation.....83
Table 7: Spillover mechanism by academic discipline.....85
Table 8: Spillover parameters estimated in included studies and related direct and total effects parameters.....87
Table 9: Studies estimating treatment-density spillovers.....91
Table 10: Search terms related to spillover effects in included texts by academic field† 92

Figure 15a: Funnel plot for spillover effects for studies estimating risk differences.....103
Figure 15b: Funnel plot for spillover effects for studies estimating risk ratios 103
Figure 15 c: Funnel plot for direct and total effects for studies estimating risk differences.....104
Figure 15 d: Funnel plot for direct and total effects for studies estimating risk ratios 104

1. Background

Description of the Problem

Interventions delivered to improve health are frequently targeted to specific populations. In many cases, such interventions benefit not only direct recipients but also those who did not receive the intervention but are connected to recipients through physical or social proximity. Such effects, which we refer to as “spillovers”, are an important component in understanding the full impact of interventions at the population-level. A “positive” spillover is in the same direction as the treatment effect, and a “negative” spillover is in the opposite direction of the treatment effect. If positive spillovers are present, studies that only estimate effects on intervention recipients will underestimate the effectiveness of the intervention. In addition, cost-effectiveness calculations that exclude such spillovers will underestimate intervention benefits. Conversely, if negative spillovers exist, evaluations only measuring intervention recipients may overestimate health impacts and cost-effectiveness.

Vaccines, one of the most efficacious and cost-effective public health interventions (Chabot et al. 2004), are a prime example of the relevance and importance of spillovers. Because of large spillovers or “herd protection” resulting from vaccination, when a critical proportion of a population is vaccinated for a particular disease, it is possible that unvaccinated individuals are also protected due to reduced transmission. Information about spillovers of other health interventions would support decisions about how best to deliver and fund large-scale interventions to improve health. In economics, the presence and magnitude of positive spillovers can justify a public subsidy for the provision of a good or service (Dybvig and Spatt 1983). As a result, studies that measure spillovers provide an evidence base that can be used to guide public funds allocation; such studies may have a larger impact than those that do not report spillover effects.

Why is it Important to do This Review

A wide range of disciplines including economics, public health, and political science discuss spillovers using a variety of terms including spillovers (Sinclair et al. 2012), interference (Cox 1958; Rosenbaum 2007; Hudgens and Halloran 2008; VanderWeele and Tchetgen Tchetgen 2011; Tchetgen and VanderWeele 2012), contamination (Hayes et al. 2000; Vermeersch and Kremer 2004), herd immunity (Fine 1993; Fox et al. 1995; John and Samuel 2000), stable unit treatment value assignment (SUTVA) violations (Rubin 1990), stability violations (Halloran and Struchiner 1995), and indirect effects (Halloran et al. 1991; VanderWeele et al. 2012). A rich literature describes empirical and causal inference methods for estimating spillovers of vaccines (Longini et al. 1988; Halloran and Struchiner 1991; Halloran and Struchiner 1995; Hudgens and Halloran 2008; Halloran et al. 2010; VanderWeele et al. 2012; Tchetgen and VanderWeele 2012). In economics, a growing literature discusses measurement of spillovers for interventions including school-based deworming (Miguel and Kremer 2004) and insecticide-treated bed nets (Bhattacharya et al. 2013). On the whole, methods for estimating spillovers have developed independently with little cross-referencing between disciplines.

In epidemiology, spillovers have often been framed as “contamination” – an undesirable problem encountered during randomized trials that can be minimized through cluster-randomization. Indeed, it is likely that in past studies, unexpected spillovers or contamination led to the evolution of the concept of spillovers as a potentially desirable feature of an intervention. In this review, we frame spillovers in a neutral light. If spillovers are found to improve health of individuals not targeted to receive programs, their consideration in public health program evaluations may yield greater population impact and cost-effectiveness. Conversely, if an intervention is found to harm individuals who were not targeted to receive it, implementers may consider redesigning the intervention. Furthermore, negative spillovers from non-intervention recipients of a program to recipients could attenuate the effects of an otherwise beneficial intervention. For example, participants in a program delivering insecticide-treated nets to some community members may still acquire mosquito-transmitted infections if their neighbors do not use nets.

Here, we provide a synthesis of methods across disciplines and present results of a systematic review on spillovers of health interventions targeting populations in low- and middle-income countries. Our goal is to unify terminology, methods, and notation across fields to encourage more consistent, transparent reporting of spillovers. Following our review of the methods and literature on spillovers, we provide recommendations for the design, analysis and reporting of studies that wish to measure spillovers.

2. Objectives

The objectives of this study were to:

1) Identify all studies with a control group that have detected the presence of or measured spillovers arising from interventions intended to improve human health
2) Summarize methods used to detect and estimate the magnitude of spillovers, as well as identification strategies and assumptions used to make causal inference
3) Highlight methodological areas where the field would benefit from further application, development, and standardization
4) Provide guidance for the application, development, and standardization of methods to estimate spillovers in impact evaluations

3. Methodological synthesis

In this section, we summarize theories of spillover mechanisms in the public health and economics literature. We then define classes of spillover parameters using the potential outcomes framework for causal inference and discuss study designs that can be used to estimate these parameters.

3.1. Theories of spillover mechanisms

The mechanism by which spillovers occur depends on the intervention and outcomes measured as well as features of the population receiving the intervention. The method of measuring spillovers and magnitude of spillover estimates depends upon the hypothesized mechanism. In public health, the literature relevant to spillover mechanisms falls into two main domains: 1) theories of infectious disease transmission and 2) mechanisms of behavior change.

Theories of disease transmission are particularly relevant to understanding spillovers of interventions targeting infectious diseases. Mathematical models are often used to describe the rate at which members of the population progress through different stages of disease and how interventions affect this rate (Ross 1915; Anderson and May 1979; Kermack and McKendrick 1991). These models often account for contact between individuals in a population. Infectious disease interventions may result in spillovers through the following mechanisms: 1) changing in the quantity of an agent or pathogen individuals are exposed to, 2) changing in the quality of the agent individuals are exposed to (e.g., a pathogen may become drug resistant), or 3) changing in their immunity to it resulting from intervention (Hayes et al. 2000). Such models have also been used to understand how diseases spread through networks (Newman 2002).

The public health literature has also drawn on a variety of theories developed by sociologists, anthropologists, and psychologists to understand how and why humans adopt certain behaviors in a social context. These include social cognitive theory (Bandura 1986), social network theory (Barnes 1954; Bott 1957; Berkman and Syme 1979; Marsden 2006), the theory of diffusion of innovations (Haider and Kreps 2004; Rogers 2010), and more recent studies within social epidemiology that use empirical data to explore how social and behavioral norms develop and exert influence on people in the same social network or environment (Berkman and Syme 1979; Oakes and Kaufman 2006; Auchincloss and Diez Roux 2008; Smith and Christakis 2008; O'Malley and Marsden 2008; Galea et al. 2010).

Traditionally, economists have explored the concept of spillovers within markets by studying how equilibrium prices affect demand of consumers and supply of firms. Methods of identifying spillover effects in other contexts were less of a focus until more recently (Manski 1993). Development economists in particular have explored how interventions may spread via learning and imitation, norm-shaping, income effects, and other mechanisms, which we discuss below (Banerjee 1992; Bikhchandani et al. 1992; Ellison and Fudenberg 1995).

In this systematic review, we have categorized included studies based on possible mechanisms of spillover. Here, we briefly describe these mechanisms, which we identified through our systematic review:

- **Geographic proximity**: Living or spending time in close proximity to individuals receiving an intervention results in improved health outcomes. This mechanism is applicable to interventions aiming to reduce infectious diseases, such as vaccination or mass drug administration, since close proximity is nearly always required for infectious disease transmission. For example, living in the same household of someone vaccinated against pertussis may reduce transmission of pertussis to unvaccinated individuals (Préziosi and Halloran 2003).
- **Social proximity**: Knowing individuals receiving an intervention may result in improved health outcomes. This mechanism is most applicable to interventions that aim to change behaviors. For example, individuals socially connected to participants in a peer intervention to reduce drug use may be likely to reduce their own drug use as well (German et al. 2012).
- **Learning/imitation**: Non-intervention recipients learn from and imitate intervention recipients, and their change in behavior can lead to improved health outcomes. This mechanism can be viewed as a type of social proximity. For instance, people who live in the same villages as individuals receiving information about child nutrition may imitate behaviors by improving feeding practices for their own children (Singh 2011).
- **Norm-shaping**: Provision of an intervention changes norms among not only intervention recipients but also non-recipients. This is distinct from learning and imitation because it is a passive process, whereas learning and imitation are an active process. This mechanism can also be viewed as a type of social proximity. For example, a conditional cash transfer program may alter norms in certain populations by requiring certain individuals to complete health screenings to receive the cash transfer (Avitabile 2012).
- **Income/substitution effect**: Provision of an intervention to some individuals in a group results in the re-allocation of resources from those individuals to others, who may benefit from additional resources. For example, if a program provides school meals to certain children in a household, more food may be available to other children through substitution (Kazianga et al. 2009). This is also referred to as the “redistribution effect”.
- **General equilibrium effects**: Typically in reference to dynamics of an economy, these effects may occur when, for example, a cash transfer provided to some individuals influences the prices, transactions, and lending behavior of other individuals within markets in the same economy. In turn, these changes in economic behaviors may affect health outcomes of other individuals (Ribas et al. 2011).

- **Relative deprivation**: The economic status of one’s peers may adversely affect an individual’s health. For example, providing a conditional cash transfer to some teenage girls may reduce the psychological well-being of other girls who do not receive a transfer (Baird et al. 2013a).

3.2. Methods for estimating spillovers

In this section, we define classes of spillover parameters using the Neyman-Rubin potential outcomes model (Rubin 1974; Holland 1986; Neyman et al. 1990) and describe study designs for estimating these parameters. The potential outcomes framework is an approach used to estimate causes and effects with statistics used in economics, public health, statistics, and political science, among other fields. Since evaluations of the impact of health interventions frequently aim to make causal inferences, our objective in this section is to summarize parameters used to estimate spillovers and to discuss identification strategies and assumptions required within a causal inference model. In the subsequent sections describing our findings from the systematic review, we organize results by the spillover classes described in this section.

3.2.1 Estimating spillovers within the potential outcomes framework

In this and the next section, we introduce spillovers within a causal inference framework by focusing on the "double-randomized" study design (also referred to as "two-stage randomization"). This design first randomizes clusters to treatment or control, then within the treatment clusters, it randomizes individuals to treatment or control. We focus on this design because it most clearly illustrates the process of spillover parameter definition and estimation, and it allows for the most straightforward causal inference. In Section 3.2.3 we provide a more detailed discussion of alternative study designs to estimate spillovers.

We denote the potential outcome for an individual allocated to treatment j in a cluster allocated to treatment i as Y_{ij}. In practice, treatment allocation at each level can either be randomized or non-randomized. In this section, we assume treatment is randomized unless stated otherwise. At the individual level, the direct effect of an intervention can be estimated by comparing potential outcomes of treated individuals in treated clusters to those of untreated individuals in control clusters (Y_{11} - Y_{00}). One can measure within-cluster spillovers by comparing potential outcomes for individuals allocated to control who reside in clusters allocated to treatment to individuals allocated to control who reside in clusters allocated to control (Y_{10} - Y_{00}).

In the next section, we define other types of spillover parameters, and we use statistical rather than counterfactual notation with the hope that it will make complicated spillover parameters intuitive to readers. For example, the within-cluster spillover effect (E[Y_{10} - Y_{00}]) can be estimated via a simple comparison of conditional means, which can be written in statistical notation as E[Y|X=1, T=0] - E[Y|X=0, T=0], where X denotes cluster treatment assignment and T denotes individual treatment assignment. If treatment allocation is randomized at both the cluster and individual level, then the within-cluster spillover causal parameter E[Y_{10}-Y_{00}] can be identified by the statistical parameter E[Y| X=1, T=0] - E[Y|X=0, T=0].

If treatment allocation is not randomized at both cluster and individual levels, other statistical estimands and assumptions are needed to identify this causal parameter. These assumptions include the "randomization assumption" (also referred to as "strong ignorability" or the "experimental treatment assignment" assumption), which is equivalent to assuming no unmeasured confounding (Rubin 1974; Rubin 1976). Another assumption is the stable unit treatment value assumption (SUTVA) or "no interference" assumption (Cox 1958), which states an individual's potential outcome is not affected by the treatment assignment of other individuals (Rubin 1990). The concepts underlying SUTVA are akin to the "reflection problem" described by

Manski, which arises when investigators try to understand how individuals' behaviors are affected by the behavior of others in their population (Manski 1993). The potential outcomes framework defined by Rubin requires that SUTVA hold. However, in studies of interventions that could spread between participants in the target population, whether by infectious disease transmission, information sharing, or mimicking of behaviors through social networks, SUTVA is violated. Thus, any study in which the investigator hypothesizes spillovers might be present inherently violates SUTVA. Halloran and others have extended the potential outcomes framework to allow for causal inference in the estimation of direct effects and within-cluster spillover effects when SUTVA is violated. They define potential outcomes that allow for an individual's counterfactual to depend not only on their treatment but also on the treatment assignment of others in the population (Halloran and Struchiner 1995; Hudgens and Halloran 2008; VanderWeele and Tchetgen Tchetgen 2011; VanderWeele et al. 2012). In section 3.2.3 we discuss study designs that can validly estimate spillovers when SUTVA does not hold.

3.2.2 Classes of spillover parameters

In this section, we define several classes of spillover parameters using unified notation from the Neyman-Rubin potential outcomes model (Rubin 1974; Holland 1986; Neyman et al. 1990). These definitions were informed by our review of the literature. We present these parameter classes prior to reporting the systematic review findings so that these parameter definitions can guide our reporting of spillover evidence.

We define six classes of spillover parameters using unified notation: 1) treatment coverage mean/effect, 2) within-cluster spillovers, 3) distance-based spillovers, 4) spillovers conditional on exposure to cases, 5) spillovers conditional on treatment density, and 6) social network spillovers. We have defined these six classes of spillover parameters because they were the most common and theoretically of the greatest interest in our literature review. However, other classes of parameters may exist. Table 1 summarizes the primary design type, scale, and mechanism for each spillover class. The scale of spillover listed in Table 1 refers to the expected magnitude of spillovers of geographic or social distances. A "small" scale spillover might only occur within households, whereas a large-scale spillover might occur within an entire village.

The choice of parameter depends on the hypothesized mechanism and scale of the spillover of interest. Some parameters may be more appropriate for spillovers through physical vs. social mechanisms. With regard to scale, certain parameters may be more appropriate for detecting spillovers on a small scale, whereas others may only be appropriate for those expected to occur over large social or physical distances.

In our discussion of each parameter below, we refer to treatment allocation in accordance with the original study design—in other words, each parameter is an intention-to-treat parameter. However, when compliance is imperfect, each parameter could also be estimated "as treated" (i.e., "per protocol") such that the treatment status of a cluster or individual is determined based on whether treatment was received regardless of the original treatment allocation. When estimating spillover parameters with imperfect compliance, the choice between an intention-to-treat or as treated analysis is subject to the usual trade offs between selection bias and underestimation of the statistical parameter (Little and Rubin 2000).

Treatment coverage effect

This parameter assesses whether greater intervention coverage is associated with a reduced risk of illness either among all individuals or among those who did not receive the intervention. The presence of such an association may provide evidence of reduced transmission of disease in areas with higher intervention coverage. When the association is measured among all individuals (both treated and untreated), it combines both direct and spillover effects. When it is estimated only among untreated individuals, it measures spillover effects. We define the "treatment coverage mean" as a parameter that compares the mean risk of illness (or other health outcome) over different levels of treatment coverage and the "treatment coverage effect" as a parameter that compares a measure of association (e.g., a difference of means) between a treatment and outcome at different levels of treatment coverage.

Treatment coverage mean: $E[Y_c \mid P_c = p_c]$ for $c = 1,...,C$ (1)

Treatment coverage effect: $E[Y_c \mid P_c = p_c] - E[Y_c \mid P_c = p_c - \delta]$ for $c = 1,...,C$ (2)

Where Y_c is the mean outcome (e.g. risk of illness) in area c, P_c is the proportion allocated to or receiving treatment in area c, and δ is a pre-defined difference in p_c. We defined the treatment coverage effect as a difference of means, but it could also be defined as a ratio of means (e.g., a relative risk). An advantage of this parameter is that it can often easily be calculated even if spillover measurement was not built into the original study design. However, it is likely that in many studies, the association between intervention coverage and risk of illness will be confounded by factors such as socioeconomic status. Furthermore, this parameter averages over groups of individuals, so at best, it allows investigators to make ecologic inferences and may be subject to the ecologic fallacy (Morgenstern 1982). Thus, it is generally preferable to adjust these estimates for potential confounders. In addition, the association between intervention coverage and illness is likely very sensitive to the definition of the area in which each measure is calculated. We discuss the implications of area definition further below.

In a randomized trial assessing cholera vaccines, Ali et al. compared the incidence of cholera among placebo recipients living in neighborhoods with varying levels of cholera vaccine coverage (Ali et al. 2013). They found that the risk of cholera decreased among placebo recipients as neighborhood vaccine coverage increased.

Within-cluster spillovers

In a double-randomized design, this parameter compares outcomes among individuals that did not receive treatment in clusters allocated to treatment to outcomes among individuals in clusters allocated to the control group (Figures 1 and 2). Figure 1 includes related parameters – the direct and total effect – that can be estimated using the same design. Direct effects compare outcomes among individuals allocated to treatment in clusters allocated to treatment to individuals allocated to control in clusters allocated to control. Total effects compare outcomes among all individuals in clusters allocated to treatment compared to outcomes among all individuals in clusters allocated to control. Frequently, within-cluster parameters condition on other variables, such as whether or not an individual was eligible for the intervention.

Within-cluster spillover effect: $E[Y \mid T=0, X=1] - E[Y \mid T=0, X=0]$ (3)

Within-cluster spillovers are an appropriate measure when spillovers are expected on a small scale. They are relatively convenient to estimate since individuals allocated to treatment and control within randomized clusters are often well defined and reachable. For this estimand to identify the causal effect, there must be a buffer zone separating treatment and control clusters of a sufficient size so that it can reasonably be assumed that the intervention does not spill over into the control clusters (Hudgens and Halloran 2008). If buffer zones are too small, this estimand will be closer to the null relative to the true causal effect. In addition, for this parameter to be unbiased, individual treatment allocation in clusters allocated to treatment must be randomized, otherwise selection bias may occur within clusters. We discuss designs for estimation of this parameter further in Section 3.2.3.

In a study of a sexual health education course in schools, Chong et al. first randomized schools to treatment or control, and then within treated schools they randomized classrooms to treatment or control (Chong et al. 2013). They hypothesized that students in control classrooms in treated schools might gain knowledge as a result of being in a school in which other classrooms received the intervention. This double-randomized design allowed for unbiased estimation of the within-cluster spillover effect of the program. They found a non-significant 7.6% decrease in knowledge scores among students in control classrooms, suggesting that within-school spillovers were not present.

<u>Distance-based spillovers</u>

When spillovers are hypothesized to occur on a larger scale, estimation of distance-based spillovers may be of interest. Investigators can compare outcomes among individuals allocated to control who reside in close proximity to the clusters allocated to treatment and control. One can define areas surrounding the clusters at fixed distances *d* and compare outcomes of individuals allocated to control near clusters allocated to treatment or control at a given distance. For example, in Figure 3, individuals allocated to control within 500 m of the treatment and control clusters could be compared. Outcomes of individuals within 1000 m in these two groups could also be compared. The distance-based spillover estimates for varying distances can be compared to assess whether the spillover effect decays as the distance from the treated cluster increases.

Distance-based spillover effect: $E[Y|T=0,X=1,D=d] - E[Y|T=0,X=0,D=d]$ (4)

For this parameter to be well defined, clusters allocated to treatment and control must be sufficiently separated so that individuals who are close to the margins of the clusters can only be influenced by the cluster they are associated with. If an individual can be influenced by the treatment assignment of multiple clusters, then it is impossible to separate the effects of the different clusters without assuming a model for how the spillovers change with distance. One potential challenge with this parameter is that in many cases, it may be logistically difficult to randomize to clusters far enough away from each other to meet this requirement. Individuals allocated to treatment may also experience distance-based spillovers if a higher density of individuals allocated to treatment in nearby areas is associated with the outcome.

In a randomized trial of an immunization campaign with incentives, Banerjee et al. estimated a distance-based spillover similar to the parameter described above. They randomized villages to receive an immunization campaign, an immunization campaign with incentives, or to a control

group (Banerjee et al. 2010). For each intervention village, they also randomly selected a village within 6 km to measure spillovers and compared outcomes to those in the control group. However, since only a single village near each treated village was measured, the design did not permit assessment of the relationship of spillovers over varying distances.

Spillovers conditional on treatment density

One can also measure the number of units allocated to treatment (N_t) or proportion of units allocated to treatment (P) within a given distance (d) of units allocated to control to assess whether the probability of the outcome is associated with the local density of treatment. For example, as shown in Figure 4, one can compare outcomes among individuals allocated to control for whom 90% of individuals within a 30 meter radius were treated to outcomes among those for whom 0% were allocated to treatment. The counterfactual group for this class of parameters can vary depending on the minimum number ($N_t(d)$) or percentage ($P(d)$) of units allocated to treatment within a specific distance d.

Spillover effect conditional on treatment density:

$$E[Y|T=0, X=1, N_t(d)=n_t(d) + \delta, N(d)=n(d)] - E[Y|T=0, X=0, N_t(d)=n_t(d), N(d)=n(d)] \quad (5a)$$

$$E[Y|T=0, X=1, P(d)=p(d)+ \delta] - E[Y|T=0, X=0, P(d)=p(d)] \quad (5b)$$

δ indicates a pre-determined difference in the treatment density. For example, if one were to compare outcomes at 90% and 10% treatment densities, δ would equal 80. Estimand 5a controls for both the number of people allocated to treatment within a certain distance ($N_t(d)$) and the number of people within a certain distance ($N(d)$) to account for differing population sizes across study areas. In some cases 5a and 5b may yield similar estimates; however, when there are very few individuals in a cluster, the proportion cannot be estimated accurately, and conditioning on the number allocated to treatment and population size is more sensible.

This parameter can be defined based on the expected heterogeneity in the proportion of individuals allocated to treatment in a community. In some cases, a narrow range of treatment proportions across distances might be expected. If individuals allocated to treatment are evenly distributed across all study areas, there might be no area with 0% individuals allocated to treatment. In this case, the counterfactual group may be defined such that $N_t(d)$ or $P(d)$ is greater than zero. On the other hand, when a wide range of treatment densities is observed, a potential dose-response pattern can be assessed by comparing the outcomes among individuals allocated to control over a range of treatment densities.

When estimating this parameter, the choice of the area in which to measure treatment density may be driven by the hypothesized scale and mechanism of spillover as well as by logistics. In the example depicted in Figure 4, density is measured within a circle with a particular individual at the center. However, other shapes are also possible, such as census tracts, and the area of measurement might not always be centered upon a particular individual. Neighborhoods may be defined around particular households or individuals using spatial clustering techniques such as centroid clustering or k-means clustering (Everitt et al. 2011). When choosing the area in which to measure treatment density, because of the problem of “modifiable nature of aerial units” described by Openshaw, a wide range of different results can be observed depending on how

an area is defined; in some cases, the definition of area units can lead to spurious findings— a form of the ecologic fallacy (Openshaw 1984).

This class of parameters can also be estimated when it is hypothesized that spillovers occur through social networks. In this case, one could condition on the proportion of social network nodes that were allocated to treatment within a given social distance metric. One could compare outcomes among individuals with a high proportion of social network connections allocated to treatment to those among individuals with no social network connections who were allocated to treatment. Thus, the distance *d* within which the proportion of individuals allocated to treatment ($P(d)$) is measured can index either physical or social distance.

Bhattacharya et al. estimated direct effects conditional on treatment density in a study in which subsidies for insecticide-treated bed nets (ITNs) were randomized to households (Bhattacharya et al. 2013). They hypothesized that the proportion of individuals who are offered subsidized ITNs in a neighborhood may affect ITN acquisition among other individuals who were not offered subsidies. While estimating the effect of the randomized subsidy on ITN purchases, to account for possible spillover effects, they estimated the fraction of households within 250 meters, 500 meters, and 1000 meters of each randomized household that were using ITNs. They found significant differences in their estimate of the effect of subsidies when they accounted for spillovers compared to when they did not, suggesting that spillovers were present in this study.

<u>Spillovers conditional on exposure to cases</u>

In the vaccine literature, there is a class of parameters used to understand interruptions in transmission resulting from a vaccination. These parameters condition on exposure to individuals who already have the outcome of interest (i.e., “cases”). They are typically estimated in studies in which cases are identified through surveillance and then their outcome-free, susceptible household members are enrolled as “controls”. For example, to understand the extent to which cholera vaccination protects susceptible individuals, one could compare outcomes among vaccinated controls living in households with cases to outcomes of unvaccinated controls living in households with cases. This parameter is called vaccine efficacy for susceptibility (VE_S) and is depicted in Panel A of Figure 5. To measure whether a vaccinated case is less likely to transmit the disease to controls than an unvaccinated individual, one can estimate the vaccine efficacy for infectiousness (VE_I), as shown in Panel B of Figure 5 (Halloran et al. 2010). Equations 6 and 7 below are written in their standard formulation in the vaccine literature – as one minus the relative risk since a protective effect is assumed. However, they could also be estimated on other scales (e.g., the additive scale).

Vaccine efficacy for susceptibility based on transmission probability:

$$VE_S = 1 - [(E[Y_i \mid T_i = 1, Y_j = 1]) / (E[Y_i \mid T_i = 0, Y_j = 1])] \quad (6)$$

Vaccine efficacy for infectiousness:

$$VE_I = 1 - [(E[Y_i \mid T_j = 1, Y_j = 1]) / (E[Y_i \mid T_j = 0, Y_j = 1])] \quad (7)$$

The index *i* indicates the treatment or outcome status for susceptible individual (Y_i=0) and *j* indicates the treatment or outcome status for the case (Y_j=1), and individuals *i* and *j* are

exposed to each other in the same cluster. Note that the only difference between equations 6 and 7 is that for vaccine efficacy for susceptibility (VE_S), the parameter conditions on the treatment status of the individual whose outcome we are measuring (T_i), whereas for vaccine efficacy for infectiousness (VE_I) the parameter conditions on the treatment status of the index case (T_j).

These parameters are akin to the within-cluster spillover parameter defined in equation 3 in that they compare outcomes between individuals allocated to control in clusters with differing treatment assignments. The key difference is that within-cluster spillover parameters condition on treatment status, whereas the VE_S and VE_I condition on outcome status as well as treatment status. In vaccine studies, these parameters are often measured in small transmission units, such as households. However, these parameters could also be estimated with larger units of clustering and potentially within social networks as well.

Préziosi and Halloran estimated vaccine efficacy for susceptibility and infectiousness in a study of the pertussis vaccine (Préziosi and Halloran 2003). They enrolled individuals who developed pertussis within an active surveillance population at a time when 77% of children under 5 years of age were vaccinated for pertussis. They enrolled children under age 15 years who lived in the same households as a pertussis case and had no history of pertussis, measured secondary cases of pertussis among contacts, and estimated the VE_S and VE_I.

<u>Social network spillovers</u>

When the purported mechanism of spillover is through social rather than physical proximity, parameters can be estimated using data on social networks. A variety of parameters can be defined for estimation within networks (Banerjee et al. 2013; Shakya et al. 2014); we focus on one type of a parameter here. Such studies typically define the initially enrolled subject as the "ego" and the person socially connected to the ego who may influence their behavior as the "alter". One can compare outcomes among alters allocated to control who are connected to egos allocated to treatment to outcomes among alters allocated to control connected to other egos allocated to control. Figure 6 shows how outcomes could be compared amongst the closest alters; it would also be possible to compare outcomes among alters at further social distances if spillovers were hypothesized to spread on a greater scale. A limitation of this class of parameters is that social network information frequently requires a near census of a target population in order to define connections between a large number of individuals, and thus data needed for such studies can be cumbersome to collect.

Spillover among social network members: $E[Y_i \mid T_i = 0, T_j = 1] - E[Y_i \mid T_i = 0, T_j = 0]$ (8)

where *i* indexes the alter and *j* indexes the ego.

German et al. estimated social network spillovers in a study of a randomized peer network intervention on depression among drug users (German et al. 2012). After randomizing drug users to receive the intervention or to a control group, they asked each participant to invite sex partners or friends who used drugs and were not already participating in the study to enroll. They compared outcomes among individuals invited by intervention vs. control group participants to assess whether the peer network intervention had spillover effects.

3.2.3 Study designs for spillover estimation

In this section, we describe study designs that can be used to estimate the parameters discussed above and summarize limitations of those methods. These include the double-randomized design, cluster-randomized design, individually randomized design, matched designs, case-control designs, regression discontinuity designs, and instrumental variable designs. All of the designs we discuss assume independent units at some level – without replicated, independent units in a study, statistical inference becomes very difficult (van der Laan 2012).

For any study design, regardless of the spillover parameter estimated, if contamination of the control group occurs, in general, spillover estimates can be considered lower bounds of the true spillover effect under certain assumptions. Specifically, one must assume that the treatment effect is of the same magnitude or less in control clusters. In addition, SUTVA must hold; otherwise, it is possible that the re-composition of treatment and control units within a cluster resulting from contamination may alter transmission dynamics and cause the treatment effect to be biased away from the null.

<u>The double-randomized design</u>

As described above, this design randomly allocates clusters and individuals to treatment or control. The proportion of individuals allocated to treatment within treatment clusters can vary depending on the research question and study population. The double-randomized design can be used to estimate within-cluster spillovers as illustrated in Figure 2 and equation 3. If there is sufficient distance between clusters and sufficient clusters with the same proportion of individuals allocated to treatment, it could be used to estimate distance-based spillovers (equation 4). In addition, this design could be used to estimate the spillover effect conditional on treatment density (equation 5) if a large enough number of clusters is assigned to different proportions of treatment. However, such a design might be very difficult to implement in practice because each level of treatment density would effectively constitute a different arm in the study, so the required number of clusters and total sample size would be very large. The vaccine efficacy for susceptibility (VE_S) can also be estimated with this design (equation 6). To estimate the vaccine efficacy for infectiousness (VE_I) using this design (equation 7), one could study the subset of clusters in which at least one individual had the outcome.

Causal inference under the double-randomized design

As discussed above, SUTVA must hold in order to make causal inferences under the potential outcomes framework, but it is typically the case that SUTVA does not hold when spillovers are likely to be present. The strength of the double-randomized design lies in its ability to estimate treatment and spillover effects when SUTVA is violated. By randomizing clusters of individuals to treatment or control, the design can validly measure the total effect of the intervention as long as the clusters remain independent. By also randomizing treatment to individuals within randomized clusters, the design can also estimate valid within-cluster spillover effects. Halloran and others have showed that unbiased spillover effects could be estimated using a double-randomized design under a set of assumptions (Halloran and Struchiner 1995; Longini et al. 1998; Hudgens and Halloran 2008; Angelucci and Maro 2010; VanderWeele and Tchetgen Tchetgen 2011). One of these assumptions is “partial interference”, which states that there must be a sufficient buffer zone between treatment and control clusters such that it is unlikely that the

interventions could affect the control group. If there are spillovers between clusters, clusters can no longer be independent, and any estimation procedures relying on the SUTVA assumption will not hold. Thus, it is best to attempt to enroll control clusters with sufficient physical or social distance from treated clusters so that independence can safely be assumed. Finally, under this design, standard errors for both direct or spillover effects must be adjusted to account for the clustering of outcomes within clusters.

Cluster-randomized designs

There are several variations of the cluster-randomized design that can be used to estimate spillovers rigorously. Within-cluster spillovers can be estimated as in the double-randomized design but with non-random allocation of treatment within treatment clusters (Clemens et al. 2011). The mini-community design is a variation of the cluster-randomized design that assigns treatment to small transmission units, such as households, and can be used to estimate spillovers (Halloran 2012). Within-cluster spillovers can also be assessed in existing cluster-randomized trials by comparing outcomes among individuals in treatment and control clusters who reside near individuals enrolled in the cluster-randomized trial (Colford 2015). In most cases, individuals near to original trial participants should be similar on average in the treatment and control clusters; as a result, this design minimizes confounding.

To measure distance-based spillovers, clusters can be randomized to treatment or control, and then additional clusters near clusters allocated to treatment can be enrolled to estimate within-cluster spillovers (Banerjee et al. 2010). This design yields high quality evidence of spillovers as long as villages enrolled to measure spillovers nearby interventions are enrolled far enough away from control clusters that investigators are confident that the chance of negative spillovers from the control clusters into the spillover clusters is minimal.

Causal inference under the cluster-randomized design

In cluster-randomized designs, it is frequently assumed that SUTVA is violated at the individual level, so investigators attempt to enroll clusters with sufficient distance between them so that clusters can be considered independent. To make causal inferences about spillovers within cluster-randomized trials, investigators must also assume that there are no systematic differences between individuals who received and did not receive the intervention in treatment clusters. In many cases, this assumption is not reasonable. For example, in a randomized promotion design (Gertler 2010), an intervention is promoted at the cluster level; which individuals choose to participate in the intervention likely depends on individual characteristics, which could result in selection bias, and thus the intention-to-treat direct effect and spillover effects may be biased. If individuals who choose to participate are likely to benefit more from an intervention than a randomly sampled group of individuals, then the intention-to-treat effects would be overestimated; conversely, if those who participate are less likely to benefit, the intention-to-treat effects would be underestimated. One approach to minimizing bias is to match untreated individuals in treatment clusters to individuals in control clusters. Doing so minimizes differences in measured confounders between untreated individuals in treated and control clusters. For example, Janssens et al. used propensity score matching to match individuals in villages that received a women's empowerment program who did not participate in the program to similar individuals in villages without the program (Janssens 2005).

When estimating direct or spillover effects with a cluster-randomized design, standard errors must be adjusted to account for the clustering of outcomes within clusters. As for the double-randomized design, to minimize SUTVA violations, investigators must also attempt to enroll control clusters sufficiently far apart from treatment clusters that there are no spillovers into the control clusters.

Individually randomized designs

Individually randomized designs can be used to estimate spillovers conditional on exposure to cases or treatment density. Investigators can estimate vaccine efficacy parameters by nesting the mini-community design within an individually randomized trial where one member of each transmission unit (e.g. household) is randomized in the trial. In this case, the study would need to also enroll household members of the randomized individuals who are incident cases (Halloran 2012). By subsetting to households with at least one case, it is possible to estimate spillover parameters conditional on exposure to cases (VE_S and VE_I).

Individually randomized designs can also estimate social network spillovers if individuals assigned to treatment and control have independent social networks. However, identifying spillovers in social networks is complicated if individuals allocated to treatment are not independent. van der Laan has described assumptions required to identify treatment effects within networks (van der Laan 2012).

Individually randomized designs can also estimate spillovers conditional on treatment density. Because the distance between randomized individuals is randomly determined, assuming high compliance, one can take advantage of random variation in the proportion of treated individuals to estimate spillover effects (Dupas 2014). A disadvantage of this approach relative to the use of a double-randomized design is that the variation in treatment densities is not fixed by design, which could limit the range of treatment densities that can be considered in the analysis.

Compared to cluster-randomized designs, individually randomized designs are more susceptible to SUTVA violations since the design does not inherently build in physical or social distance between individuals. The risk of spillovers in individually randomized designs is particularly high when studying infectious disease outcomes due to the nature of disease transmission through physical proximity. Thus, investigators using such designs to estimate spillovers must carefully assess whether the assumptions underlying their identification strategy are reasonable. Studies enrolling individuals over large physical or social distances can minimize the risk of SUTVA violations.

Observational study designs

A variety of observational study designs can be used to estimate spillover parameters including matching (Ribas et al. 2011), case-control studies (Préziosi and Halloran 2003), regression discontinuity designs (Ziegelhöfer 2012), and instrumental variable designs (Godlonton and Thornton 2012). In any of the following designs, investigators must take care to carefully assess whether units are independent (i.e., whether SUTVA was violated) and account for clustering within groups when estimating standard errors if needed.

Matched designs

Studies can utilize matching algorithms in either the design or analysis stage in order to increase comparability of the treatment and control group. Investigators can use algorithms to match clusters that did not receive an intervention to clusters that received the intervention; the same can be done to match individuals. In theory, matching of clusters and individuals could occur at either the design or analysis stage; however, in practice it is usually only feasible to match individuals at the analysis stage. To match units, multivariate matching algorithms can be used, such as propensity score matching (Rosenbaum and Rubin 1983; Rosenbaum and Rubin 1985) or genetic matching (Diamond and Sekhon 2013), if there are too many matching characteristics to enable an exact match.

A two-stage design analogous to the double-randomized design could be employed with matching such that clusters allocated to treatment are first matched to clusters allocated to control, and then individuals within clusters allocated to treatment are matched to individuals in clusters allocated to control. Such a design would allow for estimation of the same parameters that can be estimated in a double-randomized design and would minimize confounding of measured covariates.

Spillovers estimated using matched designs can only have a causal interpretation if investigators assume that treatment allocation was essentially randomized conditional on observed covariates. This is called the “strong ignorability” assumption in the Neyman-Rubin causal model (Rubin 1978) and is a universal challenge for observational studies. The magnitude of bias in parameters estimated with a matched design depends on the extent to which the matching process can approximate randomization and, as mentioned above, whether the method of recruiting individuals within clusters could result in selection bias.

Buttenheim et al. used propensity score matching in a study estimating spillovers of a school feeding program (Buttenheim et al. 2011). The program was not randomized, so they used propensity score matching to weight observations in villages with the program by 1 and in villages without the program by $p/(1-p)$ where p is the modeled probability of treatment. They hypothesized that younger and older siblings of children whose schools provide meals may also experience improved nutritional status, thus they also measured outcomes in siblings. They did not find statistically significant spillover effects.

Case-control designs

Vaccine studies frequently use case-control designs to estimate vaccine efficacy parameters (VE_I and VE_S) since the design conditions on outcome status during enrollment (equations 6 and 7). Studies typically enroll an outcome-free individual in the same household as the individual with the outcome; thus, they match on household status and treat the household as the transmission unit. By doing so, they are able to minimize confounding. Such studies usually assume that households are independent and that outcome-free individuals enrolled are only exposed to the individual with the outcome in their household (Halloran et al. 2010). For example, the study conducted by Préziosi and Halloran discussed in Section 3.2.2 enrolled households with children who developed pertussis, then enrolled children in the same households with no pertussis history as controls (Préziosi and Halloran 2003). A limitation of using a case-control design instead of a randomized design is that exposure to infected cases

might not be balanced between individuals allocated to treatment or control, which could bias estimates of vaccine efficacy (Halloran et al. 2010).

Regression discontinuity designs

Regression discontinuity designs are widely used in economics and are increasingly used in epidemiology (Imbens and Lemieux 2007; Bor et al. 2014). Such designs use a continuous variable to assign individuals to groups below and above a cutoff. For individuals with values near the cutoff, treatment assignment is ignorable and approximates randomization well under certain conditions. When such designs are used to allocate clusters to treatment, within-cluster and distance-based spillovers can be estimated. When the design is used to allocate individuals to treatment, spillover parameters conditional on exposure to cases or treatment density could be estimated using the approaches described above for individually randomized trials.

Ziegelhöfer et al. utilized a regression discontinuity design to estimate spillovers of a community water program on diarrhea (Ziegelhöfer 2012). Villages in Guinea were eligible for the program if the investment cost required to install the water infrastructure was less than 100 Euros per inhabitant at the time the program started. Since the program was not randomized, Ziegelhöfer et al. used a regression discontinuity design to compare outcomes among villages with investment just below and just above 100 Euros per inhabitant. In estimating the effect of the program, they conditioned on the proportion of individuals that received treatment within 3 km to estimate potential spillover effects. Comparing individuals' outcomes in villages just above and below the investment cost cutoff, they found that the proportion of individuals that received treatment was associated with a decreased probability of diarrhea.

When this design is used to evaluate an intervention deployed at a particular time and time is used to create a discontinuity, the design is called interrupted time series. Such an approach has been used to compare disease rates before and after introduction of a vaccine in a population. For example, do Carmo et al. compared diarrhea mortality before and after the introduction of the rotavirus vaccine in Brazil (do Carmo et al. 2011) and Grijalva et al. compared pneumonia before and after pneumococcal conjugate vaccine was introduced in the United States (Grijalva et al. 2007) using an interrupted time series analysis. To our knowledge, there have been no published studies utilizing this design to estimate spillovers of other interventions. However, such an approach could be used to rigorously estimate direct and spillover effects of a large program deployed to a population at a single time.

Instrumental variable designs

Instrumental variables are a technique used to control for confounding in observational studies (Angrist et al. 1996; Greenland 2000). An instrument is a variable that is associated with treatment status but does not directly affect the outcome; this is referred to as the exclusion restriction. If such an instrument (Z) exists, then the association between an intervention (T) and an outcome (Y) can be expressed as the ratio of the association between Z and Y and the association between Z and T. Estimation of spillover effects using instrumental variables is similar to estimation of direct or total effects. An instrument could be used to estimate both spillover, direct, or total effects. As with any analysis using instrumental variables, for estimates to be unbiased, the instrument must be independent of confounders of the intervention and

outcome, the instrument must be associated with the intervention, and the exclusion restriction must be met (Greenland 2000).

Godlonton and Thornton estimated spillovers using an instrumental variables approach. They used randomized incentives to learn HIV test results as an instrumental variable to assess whether the proportion of a person's neighbors who learned test results increased the probability they would learn their own result (Godlonton and Thornton 2012). Using a randomized treatment as an instrument allows for rigorous measurement of spillover and direct effects because it guarantees that the exclusion criterion is met.

Before and after designs

Some studies measure spillovers by comparing outcomes of untreated individuals before and after an intervention. Such designs are typically of weak quality because it is not possible to assess whether changes resulted from the intervention or from other factors varying over time. However, when an intervention is introduced to a population in which it has never been previously implemented, impact evaluations that measure outcomes that are highly specific to the intervention immediately before and after intervention can make rigorous inferences. Such a design utilizes a similar identification strategy to the regression discontinuity design; it assumes that individuals' potential outcomes are likely to be highly comparable before and after intervention as long as the timing of the intervention is not strongly associated with potential confounders. Before and after designs have been used in the vaccine literature when a new vaccine is introduced to a country and, for example, all infants born after a certain date receive the vaccine (Halloran et al. 2010; Curns et al. 2010; Hammitt et al. 2014). Spillovers can be estimated in such studies by comparing outcomes among untreated individuals before and after the introduction of the intervention in a population.

4. Methods

4.1. Protocol and registration

We attempted to register our protocol with the Campbell Coordination International Development Coordinating Group (IDCG). However, because our protocol included a synthesis of methods in addition to a systematic review, the IDCG did not accept our protocol. Instead, the International Initiative for Impact Evaluation (3ie), which funded this endeavor, supported the development of the protocol and provided both internal and external review.

4.2. Criteria for considering studies for this review

Types of Participants

Studies and interventions must have been located in a low, lower-middle, or upper-middle income countries as defined by the World Bank. They currently classify countries using 2011 gross national income per capita as follows: low income, $1,025 or less; lower-middle income, $1,026 - $4,035; upper-middle income, $4,036 - $12,475 (World Bank 2012).

Types of Interventions

We define “interventions” as the provision of services or health care through a program or study. To be included, a study must either:

1. Evaluate interventions related to health or
2. Evaluate interventions related to at least one of the following domains of human well-being and measure health outcomes: agriculture, education, employment generation, empowerment, governance (including voting and corruption), health, microfinance, migration.

Medical or hospital-based interventions that focused on noninfectious diseases with no behavior change component (e.g. hysterectomy) were not included since we did not deem it plausible for spillovers to occur from such interventions.

Spillover definition

In order to be included in the review, a study must have:

1. Measured outcomes among a group that was not targeted for intervention but that was connected geographically, socially, or by some other means to the intervention group (i.e. the “spillover group”)
2. Clearly articulated a counterfactual for the spillover group that was measured or estimated in the study

In determining eligibility, we did not distinguish between whether individuals targeted for intervention complied with their treatment allocation or not. Some studies did not state whether individuals who did not receive the intervention were targeted or not; in other words, the individuals who did not receive treatment may have included a mix of eligible, non-compliers

and ineligible individuals. In such cases, we included the studies if they met all other inclusion criteria.

Studies estimating either negative or positive spillovers were eligible for inclusion.

Spillovers focused on transmission to partners (e.g. male circumcision and female partner HIV acquisition) were not included in this study because we deemed both individuals to be targeted by the intervention.

In our protocol, we stated that we would only include studies that clearly articulated a hypothesized mechanism for spillover effects (e.g. geographic or social proximity). We found that this was infrequently mentioned and would have ruled out a number of otherwise eligible studies. Thus, we ultimately included studies that did not meet this criterion.

Types of Comparisons

Papers must define a comparison group that approximates the counterfactual for both the estimation of direct effects, the effect of the intervention on those who received it, and spillover effects. Authors must describe the quality of the comparison group used to estimate spillovers.

Types of Outcomes

Outcomes measured for either direct or spillover effects must be both related to human health and measured among humans or groups of humans. Theories of change for interventions to improve human health may include many intermediate steps, and evaluation of such interventions may focus on not only humans, but also other organisms or other entities along the causal pathway. We included intermediate outcomes, such as handwashing behavior. Population-based vaccine studies focusing on immunology with no clear measurement of health outcomes or diseases were not included. We did not consider food consumption, cost-effectiveness, willingness to pay for a health intervention, or consumption of health insurance to be health outcomes.

Types of Studies

We included designs to evaluate interventions that meet the following criteria:

1. Include a comparison group which is constructed in a way such that counterfactual outcomes may be estimated for both direct and spillover effects
2. Included sufficient detail about the design and comparison group to determine whether there are serious threats to internal or external validity.
3. Utilized quantitative rather than qualitative design and analysis methods.

We chose to include studies with a weak choice of comparison group or counterfactual in order to be as inclusive as possible in discussing methods used to estimate spillovers.

Other Inclusion/ Exclusion Criteria

Books were not eligible for the review. We excluded any manuscripts retrieved that were marked as drafts not to be cited.

4.3. Search Methods for Identification of Studies

Search Terms

A detailed description of our search strategy is listed in Appendix 1. We developed this strategy with substantial input from an Information Specialist at 3ie. We searched reference lists of texts classified as eligible in the original search. We also identified records that cited included texts from original search using Google Scholar. Following the search process, all records were merged, duplicates were removed, and a unique ID was assigned to each record.

Electronic Searches

We searched the electronic databases listed in Appendix 2 for articles published before 2014.

Other Searches

We hand searched all titles from 2010-2013 in the following journals: *Health Economics*, *The Journal of Development Effectiveness*, *The Lancet*, *PLoS Medicine*, and the *World Bank Policy Research Working Papers*. We chose these journals because we felt they were likely to publish results of impact evaluations of health interventions that may include spillover measurement.

Reference Management

We used Zotero and Jabref to manage references.

4.4. Data Collection and Management

Study selection

Each record retrieved was reviewed by at least one team member for relevance. Titles that were clearly not eligible for the review received no further review (e.g., those focusing on animals). Each available abstract that passed the title review was then reviewed for relevance. If an abstract was not available but a full text was, the full text was reviewed instead. Of the abstracts deemed relevant, each full text was reviewed for relevance. Team members recorded the first reason for exclusion identified for records that were deemed not to be relevant. If multiple versions of a paper were available, we included the most recent version of the paper.

Data collection process

Each included text was extracted and then checked by a second independent team member. In one case, spillover results were mentioned, and disaggregated results were not listed in the publication, but the authors mentioned that results were available upon request (Buttenheim et al. 2011). We contacted the authors to request these results but did not receive a reply.

Data items

For each included text, we extracted information about the interventions, outcomes measured, study site, primary study design, study design used to estimate spillovers, purported spillover mechanism, scale of spillover (e.g. household versus village), and whether or not spillover measurement was pre-specified. For studies that estimated within-cluster spillovers, we recorded the percent of individuals that received treatment in clusters allocated to treatment on

average. We considered spillover estimation to be pre-specified if spillover estimation was built into the original study protocol or the publication presenting original study findings. For spillover mechanisms, we recorded the purported mechanism if mentioned by the authors, and if no mechanism was mentioned, we selected possible mechanisms based on the intervention, outcome, and method of estimating spillovers. We extracted items that we pre-specified in our study protocol. For the possible mechanisms of spillover, we developed a list of potential mechanisms during our review of the included texts.

We extracted the direct effect, total effect, and spillover effects of the intervention, if reported. If multiple effects or model specifications were used to estimate the direct or total effect, we attempted to extract information that would allow the greatest comparability of the direct effect to the spillover estimates. If multiple effects were estimated, we chose the estimate that appeared to be the primary finding reported by the author.

We extracted spillover effects reported numerically in tables or text. We did not extract spillover results reported in graphical form with no numerical labels. For all types of parameters, we noted the type of measure estimated (e.g. probability difference, odds ratio, etc.), the standard error, 95% confidence interval, and any statistical tests related to the measure (e.g., t-statistic). We recorded the units of each result and whether each was adjusted for potential confounders or not.

4.5. Critical appraisal and Risk of Bias assessment

Risk of Bias Assessment

We reviewed each study for the risk of bias using criteria compiled from various fields in order to accommodate the range of studies that we will review: the Cochrane Collaboration Handbook (Higgins and Greene 2011), the Coalition for Evidence Based Policy (2010) tool for reviewing randomized controlled trials (Coalition for Evidence-Based Policy 2010), the Effective Practice and Organisation of Care Group tool for assessing bias in both randomized and non-randomized designs (Effective Practice and Organisation of Care (EPOC) Group 2009), Impact Evaluation in Practice (Gertler 2010), and “Regression Discontinuity Designs in Economics” (Lee and Lemieux 2010). These criteria are listed in Appendix 3. For each criterion, we classified a study as “yes”, “uncertain”, “no”, or not “not applicable”. Duplicate assessment of risk of bias was performed for a 20% subsample. Classification was not blinded. Co-authors of this systematic review who authored included studies did not participate in the classification of risk of bias criteria for any included studies.

For included studies that performed secondary analyses, we attempted to obtain the original publication and incorporated information from the original publication(s) into our risk of bias assessment. If a study also estimated parameters for outcomes not related to health, we only assessed risk of bias for the elements of the study that estimated effects on health outcomes.

Reporting norms vary across disciplines, and certain items were less likely to be reported in some disciplines. As a result, we felt that using a quantitative measure of risk of bias for each study summarizing the above criteria would not be a fair assessment of each study’s overall risk of bias. Instead, we augmented the Cochrane GRADE approach (Guyatt et al. 2008) to assess

the quality of evidence specific to spillover estimation (Table 2a). We developed criteria that weigh the classifications of the individual risk of bias criteria in Appendix 3 with the overall study design and quality of reporting (Balshem et al. 2011). Our rationale for these criteria is provided in Appendix 4. Then, we modified the list of factors that may increase or decrease the quality level of a body of evidence (Table 2b) (Balshem et al. 2011). We developed these criteria through an iterative process in which we revised our classification system after the initial risk of bias assessment for each study and discussion with multiple reviewers. We then classified each study's overall quality of evidence as "very low", "low", "medium", or "high".

Comprehensive validity assessment

For each included study, we assessed whether authors discussed the external validity of their findings and the representativeness of their study sample (Appendix 5). We also assessed construct validity in studies that did not measure terminal outcomes by reviewing whether intermediate outcomes measured were likely to be strongly correlated with the terminal outcome of interest.

4.6. Unit of Analysis Issues

We assessed whether studies using clustered designs accounted for clustering in the estimation of standard errors and other measures of precision (e.g. 95% confidence intervals).

4.7. Assessment of study dependence

We assessed whether common study features, including study site, intervention program, study population, and study investigators, may have led to dependence of findings across included studies.

4.8. Synthesis of findings

Data Synthesis

We identified the spillover parameter(s) estimated in each paper and compared estimates within the five parameter classes defined in Section 3.2.2. If a study estimated multiple spillover parameters, we recorded each estimate. When outcomes were measured at repeated time points, we recorded measures of spillover at each time point. Within each class of spillover parameter, we standardized treatment effects as much as possible using the available information in included studies. For binary outcomes, we calculated the percent reduction in outcomes attributable to the intervention [(1 - relative risk) x 100%]. For results presented on the additive scale, we divided by the mean of the outcome in the control group to estimate the percent reduction attributable to intervention. To generate forest plots comparing results within parameter classes, we converted parameters on the additive scale to the relative scale by dividing by the mean of the outcome among individuals not receiving treatment if such information was reported.

The information needed to convert standard errors from the additive to the relative scale was not available in most included studies; specifically, the probability of the outcome in each treatment group and the number of people in each treatment group are required. In plots

comparing estimates across studies, we presented 95% confidence intervals for the studies for which standard errors were reported or could be estimated on the relative scale. When possible, we used adjusted effect measures. Since there were very few results within a given parameter class with continuous outcomes, we did not synthesize results for such outcomes. We excluded studies with an overall high risk of bias from plots comparing results across studies.

The included texts span a wide range of interventions and outcomes, and thus, there was heterogeneity in estimates of spillovers. We do not consider it reasonable to assume that the studies included are independent and that a common treatment effect exists across all included studies (Berk and Freedman 2010). Potential sources of dependence between studies include common authors, studies evaluating the same intervention (e.g. cholera vaccines), studies conducted in the same population, and studies conducting secondary analyses upon the same primary dataset. In plots comparing results across studies, we indicate potential sources of dependence. We assessed whether meta-analysis could be conducted in subgroups (e.g. by intervention type) but concluded that even in the largest of subgroups (vaccines), there were so many different types of parameters estimated that within a given parameter class, there would be too few estimates to allow for meaningful meta-analysis.

Assessment of Reporting Bias

To assess possible reporting or publication bias, we produced funnel plots. The majority of studies did not provide sufficient information to standardize measures onto a single plot; thus, we produced separate plots for studies estimating risk ratios (or 1-RR) and risk differences. We also generated separate plots for spillover vs. total and direct effect estimates. Funnel plots only include studies that estimated effects for binary outcomes. We did not produce funnel plots for estimates using continuous outcomes because the number of different outcomes measured would not have allowed for comparison across a useful number of studies.

We also compared the proportion of statistically significant results between studies that provided documentation that the spillover effects analysis was pre-specified versus those without. Our rationale was that if the proportion of significant results was higher among studies that did not pre-specify spillovers estimation, publication bias may be present due to investigators selectively publishing statistically significant results. We assessed this for studies that included measures of statistical significance for spillover estimates (p-values, 95% confidence intervals, standard errors, and t-statistics). For treatment coverage mean and treatment coverage effect parameters, if a measure of statistical significance was available for multiple levels of coverage, we included all measures in the summary.

4.9. Additional analyses

We searched each included text for terms commonly used to describe spillovers and noted whether the terms appeared in each text.

5. Results

5.1. Description of Studies

Results of the search

We retrieved 49,749 records through our search process (Figure 7). This includes records identified by searching electronic databases and by hand searching. 35,159 records remained following duplicate removal. We dropped 3,537 records from non-bibliographic sources due to concerns about the quality of the searches yielded from these databases, many of which did not allow for standard database search techniques such as wild cards and Boolean operators. We screened 31,622 titles for relevance and determined that 11,839 were relevant. We obtained abstracts for these titles and reviewed them. Of these, 556 abstracts were deemed relevant. We obtained full texts for these records and reviewed them. Of these, 28 met study inclusion criteria.

We searched the reference list of the 28 original included texts and identified another 798 records requiring review. We performed title, abstract, and full text review on these records as described above and identified 1 additional full text from this pool that met inclusion criteria. We then used Google Scholar to identify records that cited the 28 original included texts and identified 1,766 unique records that required review. Following title, abstract, and full text review, we identified 25 additional texts that met inclusion criteria. Of these studies, 14 cited the Miguel & Kremer 2004 paper on externalities of school deworming in Kenya (Miguel and Kremer 2004; Chaudhuri 2005; Dupas 2006; Kremer and Miguel 2007; Björkman and Svensson 2009; Zivin et al. 2009; Ribas et al. 2011; Tontarawongsa et al. 2011; Fitzsimons et al. 2012; Godlonton and Thornton 2012; Ziegelhöfer 2012; Baird et al. 2013a; Joshi and Schultz 2013; Baird et al. 2013b; Ozier 2014), and of those, three studies built upon that original trial (Kremer and Miguel 2007; Baird et al. 2013b; Ozier 2014). A total of 54 records were included in this systematic review. Their characteristics are listed in Table 3.

Reasons for exclusion of full texts are listed in Appendix 6, and reasons for exclusion of each record are listed in a supplementary spreadsheet.

We extracted data from 51 studies. We could not extract data for two studies that only reported spillover effects graphically (Paul et al. 1962; Shekhawat et al. 2014) or for one study which did not provide numerical results for spillover estimates (Buttenheim et al. 2011).

Included studies

<u>Study characteristics</u>

Characteristics of each included study are listed in Table 3. Studies were conducted in 17 countries; the most common countries were Bangladesh (n=11), Kenya (n=9), and India (n=6). We noted the primary academic discipline of each included study; 25 included studies were in Economics, 25 in Public Health, and 4 in Geography. The relatively large proportion published in the economics literature likely stems from the influential study by Miguel and Kremer, which found spillover effects of school-based deworming in Kenya (Miguel and Kremer 2004).

Study designs

A wide range of designs was used in studies in the review including 38 studies that randomized treatment or studied a previously randomized population and 16 observational studies (Table 4). The ratio of randomized to observational designs was similar between studies in economics and public health. The most common design found in the included studies was a cluster-randomized trial (n=13 studies) followed by a re-analysis of a cluster-randomized trial (n=9) and a re-analysis of an individually randomized trial (n=7). Re-analyses evaluated a previously randomized intervention and utilized the original trial data and/or new sources of data in the randomized population to estimate spillovers. Sixteen studies utilized observational designs of many different types including case-control studies, cohort studies, matched studies, regression discontinuity studies, and instrumental variables analyses.

Interventions and outcomes

The most common interventions studied across academic disciplines were vaccines (n=22 studies) followed by mass drug administration for infectious disease control (n=7) and health education (n=5) (Table 5). The public health studies evaluated few other interventions, and the geography studies only evaluated vaccines. The economics studies evaluated numerous different interventions including cash transfers, empowerment programs, HIV/AIDS-related interventions, and maternal and child health interventions. A wide range of outcomes were also studied including disease outcomes such as cholera (n=9), trachoma (n=4), and pertussis (n=3) and health behavior outcomes such as screening for illness (n=1), health care visits (n=1), and voluntary counseling and testing for HIV (n=1).

To explore whether the level of treatment allocation (individual vs. cluster) was chosen in order to measure spillovers, we classified each study's rationale for treatment allocation level. For many studies, no rationale was given (n=16; Table 6). In fifteen studies, the level of treatment allocation was determined based on the level of intervention allocation. In eight out of 54 included studies, investigators explicitly stated that treatment was allocated to clusters in order to measure spillovers.

There were several programs that were commonly evaluated for spillovers in the included studies: the maternal and child health program in Matlab, Bangladesh (Chaudhuri 2005; Joshi and Schultz 2013); the PROGRESA program, which offered conditional cash transfers in Mexico (Handa et al. 2001; Avitabile 2012); cholera vaccines provided in Matlab Bangladesh (Ali et al. 2005; Emch et al. 2006; Ali et al. 2008; Emch et al. 2009; Root et al. 2011; Perez-Heydrich et al. 2014), and the Primary School Deworming Program in Busia, Kenya (Miguel and Kremer 2004; Kremer and Miguel 2007; Baird et al. 2013b; Ozier 2014). For all of these programs, the assessment of spillovers was not pre-specified and was not incorporated into initial impact evaluations of these programs. The frequent measurement of spillovers of these programs may have been a result of study designs that permitted relatively easy spillover estimation: programs were offered to particular clusters (villages or schools) and not in others, and then within clusters where the program was offered, some individuals did not participate in the program. This design allowed for the estimation of within-cluster spillovers, and in several

cases additional measurement was done that allowed for other spillover parameters to be estimated as well.

Spillover mechanisms

We classified spillover mechanisms into 9 different categories, which are listed in Table 7. 72% of the 54 included studies focused on spillovers related to geographic proximity, 31% focused on social proximity, 28% focused on learning and imitation, and 28% focused on norm-shaping. Certain mechanisms were only relevant to studies in economics: income/substitution effect, public good effect, general equilibrium effects, and relative deprivation. Geographic proximity was the sole mechanism evaluated in studies estimating within-cluster spillovers and vaccine efficacy, and other mechanisms were explored using a range of different spillover parameters.

Excluded studies

We excluded 647 full texts identified through our search process. As mentioned above, there were often multiple reasons for exclusion possible, and we only recorded the first one we encountered during screening. The most common reason for exclusion was that no spillovers were measured (n=429). Other common reasons were an invalid design, such as a cross-sectional survey with no control group or underlying identification strategy (n=51), no measurement of health outcomes (n=36), and a study design that was not empirical (e.g., a mathematical modeling study) (n=33). Fifty-one papers were excluded because the design was not valid. For example, Quian et al. conducted a study evaluating the universal varicella vaccine in Uruguay; they compared the proportion of hospitalizations due to varicella during periods of time before and after the introduction of universal vaccination (Quian et al. 2008). Their design did not include a comparison group that could serve as a counterfactual for both direct and spillover effects, thus the study was excluded. Studies that conducted cross-sectional surveys that were not evaluating a population already randomized or assigned to intervention and control groups in some way were also excluded (Perkins et al. 2007; Wamai et al. 2012).

5.2. Methodological Quality and Risk of Bias in Included Studies

Risk of bias within studies

Appendix 3 lists the percentage of studies classified as “yes”, “uncertain”, or “no” and the number of studies for which each criterion was assessed, and Appendix 7 lists the overall quality of evidence for each study. The classification for each criterion for individual studies is available as a supplementary spreadsheet.

In our assessment of individual risk of bias criteria for each study, we found that overall, most studies met general criteria for internal validity that were not specific to particular study designs. For example, in 76% of applicable studies (N=49), characteristics between intervention and control groups were similar at baseline. Only 34% of randomized studies blinded treatment assignment, 32% blinded outcome assessors, and 46% concealed allocation adequately. In many studies, it may have been logistically impossible to blind participants and outcome assessors to treatment status; nevertheless, lack of blinding remains a risk of bias regardless of logistical concerns. We found that 83% (N=42) of studies that measured outcomes within

clusters accounted for such clustering by estimating robust standard errors or using other appropriate methods (e.g. generalized estimating equations) (Appendix 3).

Our risk of bias assessment included three criteria specific to spillover estimation. In 86% (N=42) of studies, subjects included in estimation of direct effects were comparable to subjects included in estimation of spillover effects. This criterion helps assess the extent to which the total effect of an intervention may be decomposed into direct and spillover effect. Our findings indicate that in most cases, direct and spillover effects were comparable. In nearly all studies (89%; N=54), authors described the extent to which the comparison group used to estimate spillovers could be considered a valid counterfactual. While very few studies made reference to the potential outcomes framework or explicitly defined the spillover parameter estimated, many did clearly define the comparison group used in the spillover parameter estimated. For the estimation of direct effects, there was evidence of minimal contamination of the control group in only 19% of 37 studies for which this criterion was assessed. In our protocol we stated that contamination was not assessed for estimation of spillover parameters; however, in studies estimating within-cluster spillover effects and related parameters, if contamination was present, it nearly always influenced spillover effects in addition to direct or total effects. This criterion helps us assess whether possible contamination of the control group could have resulted in bias towards the null. Most studies measuring within-cluster spillovers did not mention buffer zones, which made it difficult to assess the possibility of contamination into the control group.

In our assessment of the overall quality of study evidence using the modified Cochrane GRADE criteria (Guyatt et al. 2008), 6 of the 54 included studies (11%) had high quality evidence, 30 (56%) had moderate quality, 12 (22%) had low quality, and 6 (11%) had very low quality evidence. Studies with high quality evidence utilized cluster-randomized (Hawley et al. 2003; House et al. 2009; Banerjee et al. 2010; Egere et al. 2012; Perez-Heydrich et al. 2014) and case-control designs (Préziosi and Halloran 2003). A particularly high quality design that we recommend for future studies is a cluster-randomized trial in which a village within 6 km of each village assigned to treatment was enrolled in order to measure spillovers to nearby areas (Banerjee et al. 2010). There were two studies that utilized double-randomized designs to estimate within-cluster spillovers that were initially given a high quality rating but were downgraded due to the use of self-reported or subjective outcomes, among other concerns (Baird et al. 2013a; Chong et al. 2013).

Of the studies with moderate quality of evidence, seven estimated the treatment coverage mean or treatment coverage effect (Ali et al. 2005; Emch et al. 2006; Ali et al. 2008; Emch et al. 2009; Huq et al. 2010; Root et al. 2011; Root et al. 2014) and four studies estimated spillover parameters that conditioned on distance (e.g., treatment or outcome density within nearby areas) (Kremer and Miguel 2007; Björkman and Svensson 2009; Godlonton and Thornton 2012; Ziegelhöfer 2012). Many of the low quality studies utilized observational study designs with poor control for confounding. For example, one study used an instrumental variables approach, but the instrument likely did not meet the exclusion restriction criterion, and the strength of the association between the instrument and intervention was questionable (Janssens 2005). One study was classified as having very low quality evidence because of the magnitude of spillovers relative to the direct/total effects, heterogeneity of findings in subgroup analyses, and use of subjective outcome measurement (Baird et al. 2013a). Two of the studies with low quality were

published before 1980, which may reflect the lack of standardized reporting requirements in scientific journals historically (Paul et al. 1962; Cooper and Fitch 1983).

The proportion of studies with low or very low quality evidence was similar in studies that incorporated spillover measurement into the original design (35%) compared to those which did not pre-specify spillover estimation (36%). The distribution of study quality was nearly the same when stratifying by whether or not an article was peer reviewed for very low, low, and moderate quality evidence. All high quality studies were peer reviewed. Among economics studies, the quality of evidence was very low in 16%, low in 20%, moderate in 60%, and high in 4% of studies. Among public health studies, overall quality of evidence was very low in 8%, low in 28%, moderate in 44%, and high in 20% of studies.

Study validity assessment

We also assessed studies' external validity and construct validity (Appendix 4). In general, due to limited reporting, it was difficult to assess external validity of included studies. Eighty percent of authors of included studies indicated whether external validity was likely to be high or commented on the representativeness of the study sample. Many studies conducted secondary analyses of existing datasets, and it is often the case that papers summarizing such analyses do not provide information needed to assess external validity.

In the majority of included studies, the outcomes measured were likely to be highly correlated with the outcomes that interventions were intended to impact. In addition, in 90% of studies, intermediate outcomes measured were clearly connected with terminal outcomes and the period of time in which the study was conducted was sufficient to meaningfully assess intervention impacts. Only 44% of studies discussed measurement error.

5.3. Synthesis of Results

5.3.1 Results within spillover parameter categories

In this section, we summarize findings within categories of spillover parameters. In section 5.3.2, we summarize findings within intervention categories. Twenty-two different parameters were estimated in the included studies. These are summarized in Table 8. In this section, we discuss the parameters estimated in each of the five spillover parameter classes defined in Section 3.2.2 above. Figure 8 shows the distribution of parameters estimated by discipline. Within-cluster spillover effects and the treatment coverage mean were the most commonly estimated parameters. Of the 54 included papers, one paper did not discuss or estimate spillovers, but we were able to estimate within-cluster spillover using information presented in the paper and its appendix (Azad et al. 2010).

Treatment coverage mean and treatment coverage effect

Seven papers estimated the treatment coverage mean (Table 8, parameter 1) (Cooper and Fitch 1983; Forleo-Neto et al. 1999; Ali et al. 2005; Emch et al. 2009; Huq et al. 2010; Khatib et al. 2012; Ali et al. 2013), and of these, six were in the public health literature. Eight papers estimated the treatment coverage effect (Table 8, parameter 2) by comparing measures of the efficacy of an intervention across levels of treatment coverage (Ali et al. 2005; Emch et al. 2006;

Ali et al. 2008; Emch et al. 2009; Root et al. 2011; Ali et al. 2013; Chen et al. 2014; Root et al. 2014). Fourteen of these studies estimated coverage within members of geographically defined areas around each individual, and one study estimated coverage among individuals within social networks for each individual (Root et al. 2011). Of the 13 studies measuring the treatment coverage mean or effect, 5 had low or very low quality evidence and 8 had moderate quality evidence.

Four studies estimating the treatment coverage mean focused on cholera vaccination coverage and cholera risk, allowing for comparison of results. However, these findings are likely to be highly dependent since the studies in Bangladesh utilize the same dataset. Cholera risk per 1,000 people among unvaccinated individuals declined markedly as vaccination coverage increased (Figure 9, Panel A), suggesting strong spillover effects due to herd protection. There was no decrease in risk among vaccinated individuals as vaccination coverage increases (Figure 9, Panel B), which is to be expected since those individuals were protected from illness directly through vaccination. These findings align with the theory of herd protection. Studies estimating the treatment coverage mean for other interventions and outcomes revealed a similar pattern (Cooper and Fitch 1983; Forleo-Neto et al. 1999; Huq et al. 2010).

We also compared the protective efficacy ([1-relative risk] x 100%) and vaccination coverage for three studies for which comparable data were available (Figure 10). As vaccine coverage increased, vaccine efficacy decreased. These vaccine efficacy measures are total effects – they capture both the direct effect of the vaccine and any spillover effects. The protective efficacy decreases at higher levels of vaccine coverage because as vaccine coverage increases, the risk among unvaccinated decreases, as shown in Figure 9. Thus, the difference in risk between the vaccinated and unvaccinated individuals decreases as vaccination coverage increases, the relative risk is closer to the null, and the protective efficacy is closer to zero. The two other studies that compared vaccination coverage and protective efficacy but were not included in Figure 10 found similar patterns for the *Haemophilus influenzae* type B-diptheria-tetanus-pertussis (Hib DTP) vaccine (Chen et al. 2014) and the cholera vaccine (Emch et al. 2009; Root et al. 2011).

The advantages of estimating the treatment coverage parameter are that it is easy to calculate and interpret and that non-linear trends can be detected across levels of intervention coverage. The disadvantage is that associations are likely sensitive to the way coverage levels are categorized and to the size of area in which coverage is calculated. Quintiles of the distribution were commonly used to categorize vaccination coverage (Emch et al. 2009; Khatib et al. 2012; Ali et al. 2013), which ensures the number of units will be roughly equal across coverage categories. However, since the distribution of vaccination coverage differs between studies, using percentiles can make it difficult to compare results across studies. The area within which treatment coverage was calculated varied substantially; in some studies, coverage was calculated within a cluster of several households (Ali et al. 2005), a neighborhood with radius 250m (Ali et al. 2013), or at the city level (Forleo-Neto et al. 1999; Haile et al. 2013). As discussed above, results can vary widely depending on how an area is defined (Openshaw 1984). Only some studies explicitly provided a rationale for area definition. For example, Khatib et al. used an algorithm to define a radius of equal size around each study participant using Hartley's test of homogeneity of variance (Khatib et al. 2012). The choice of area must be made

carefully, weighing the expected transmission dynamics within a given distance, the expected magnitude of spillovers, and the variability of coverage expected within a given area.

In summary, for the reasons discussed in section 3.2.2 and Appendix 7, studies estimating the treatment coverage mean or effect can only produce evidence of low or moderate quality because they assess ecologic associations. Parameters discussed in the subsequent sections provide more rigorous evidence of spillovers. Nevertheless, these parameters can be useful for hypothesis generation for new future interventions for which spillovers have yet to be measured when individual-level data is not available.

Within-cluster spillovers

Within this class of spillover parameters, two types of parameters were commonly estimated: 1) within-cluster spillovers which compared outcomes among individuals not receiving treatment in the group assigned to treatment to those among individuals not receiving treatment in the group assigned to control regardless of eligibility status (Table 8, parameter 4) (Miguel and Kremer 2004; Chidambaram et al. 2004; Janssens 2005; Sur et al. 2009; Zivin et al. 2009; Azad et al. 2010; Singh 2011; Baird et al. 2013a; Ali et al. 2013), and 2) within-cluster spillovers among individuals who were not eligible to receive treatment (Table 8, parameter 8) (Handa et al. 2001; Chaudhuri 2005; Kazianga et al. 2009; House et al. 2009; Ribas et al. 2011; Fitzsimons et al. 2012; Avitabile 2012; Contreras and Maitra 2013). For the latter parameter, eligibility was frequently based on age (e.g., for mass treatment to control trachoma infection) or income level (e.g., for enrollment in a conditional cash transfer program). Of the 22 studies estimating within-cluster spillover parameters, two had high quality evidence, 12 had moderate quality evidence, and 9 had low or very low quality evidence.

Within-cluster spillovers

Figure 11 plots within-cluster spillover effects for the four studies that reported sufficient information to allow for comparison. These spillover effects can be interpreted as the percent increase in health-promoting outcomes (e.g. child growth) and the percent decrease in adverse outcomes (e.g. cholera incidence). The strongest within-cluster spillover effects were found for a study of the typhoid vaccine in India (44% protective efficacy) (Sur et al. 2009); however, a similar study conducted in Pakistan did not find evidence of within-cluster spillovers (Khan et al. 2012). Three of the four studies included in Figure 11 were conducted in India, and thus, their results may be dependent due to shared cultural and geographic context. However given that India is a large and diverse country and that these studies evaluated different interventions, we consider this potential source of dependence to be minimal.

Figure 12 shows the standardized within-cluster spillover estimates by the level of treatment coverage within clusters receiving treatment. While there were only four studies with comparable data, the evidence suggests that spillovers were larger in studies with a higher proportion of treatment within treated clusters. Among the two studies estimating within-cluster spillovers of the typhoid vaccine, the study in which approximately 60% of people per cluster received treatment found statistically significant, large spillover effects (Sur et al. 2009), whereas the study with under 40% vaccine coverage per cluster did not find spillover effects (Khan et al. 2012).

Eight studies estimated this parameter but did not include sufficient information to allow for comparison. For example, the well-known study of a deworming program in Kenya conducted by Miguel and Kremer estimated within-school spillovers. The study reported a 12% reduction in moderate to heavy worm infections among children who attended schools in the program but did not receive deworming compared to those in control schools. A replication of this study and alternative analysis of the original data were conducted (Davey et al. 2014; Aiken et al. 2015); the replication clarified some of the design features and reporting by Miguel and Kremer and recommended presenting some alternative results. Some of the replicators' critiques were that treatment allocation was not strictly randomized and that missing data was not completely presented (Aiken et al. 2015). The replicators corrected an error in the coding of a population density variable included in the statistical model used to estimate within-cluster spillovers. Following this correction, the estimate of within-cluster spillovers was 18% instead of 12%; both estimates were statistically significant at the alpha=0.1 level. We refer readers to the thorough discussion of the replication study's findings for more details (Davey et al. 2014; Hicks et al. 2014; Aiken et al. 2015).

Within-cluster spillovers conditional on eligibility

Figure 13 plots the standardized estimates of within-cluster spillovers conditional on eligibility. This figure only includes evidence from three studies due to limited availability of information needed to convert results to a single scale, thus we can only draw limited conclusions from these findings. The strongest spillover effects were recorded for the pneumococcal vaccine, which resulted in 70% (95% CI 61, 77) reduction in vaccine-type pneumococcus among newborns born in villages where all individuals received the vaccine compared to villages where only young children received the vaccine (Egere et al. 2012). Mass azithromycin treatment was found to have strong spillover effects, resulting in a 35% decrease in trachoma among individuals who were ineligible to receive azithromycin (House et al. 2009). Provision of infant nutrition and health information did not result in any spillovers among siblings of targeted children (Fitzsimons et al. 2012).

We also compared spillover estimates and treatment coverage from these three studies (Figure 14). As for within-cluster spillovers, we found that spillover effects were larger in studies with a higher proportion of treatment within treated clusters. The largest spillovers were present for a study evaluating spillovers of the pneumococcal vaccine in villages where 100% of individuals were vaccinated (Egere et al. 2012), and the smallest spillovers were present in a study evaluating an education program covering approximately 60% of the population (Fitzsimons et al. 2012).

Other within-cluster spillover parameters

One study used propensity-score matching to match individuals who did not receive treatment in clusters assigned to treatment and control (Janssens 2005). When it is not possible to use a double-randomized design to estimate this parameter, this matching approach can minimize

bias by increasing the comparability of untreated individuals in treatment clusters and individuals in control clusters. Within-cluster spillovers can also condition on exposure to the intervention within clusters allocated to treatment (Table 8, parameter 11) (Azad et al. 2010; Baird et al. 2013a). For example, in an evaluation of a women's group intervention, investigators measured outcomes among people in intervention clusters that did not participate in the intervention but who had heard of the intervention (Azad et al. 2010).

Comparison of total and spillover effect magnitude

In a double-randomized trial, the total effect of an intervention can be decomposed into the direct effect and within-cluster spillover effect at either the individual or population level (Hudgens and Halloran 2008). For studies with available data that allowed for comparison of spillovers and direct effects, we divided the within-cluster spillover effect by the total effect to estimate the proportion of the total effect attributable to spillovers. We restricted this analysis to studies with statistically significant spillover estimates. In studies estimating the within-cluster spillover effect among ineligibles, the total effect was often estimated among eligibles, in which case it was not possible to calculate the proportion of the overall total effect (among eligibles and ineligibles) attributable to the spillover effect. Only one study that estimated within-cluster spillovers met these criteria. In a cluster-randomized study of the effect of the typhoid vaccine on typhoid, the proportion of the total effect attributable to spillovers was 72% (Sur et al. 2009).

In summary, only one study estimating within-cluster spillovers produced high quality evidence. However, this parameter can be rigorously estimated relatively easily within cluster-randomized trials or double-randomized trials. A common potential source of bias for these parameters in cluster-randomized trials is selection bias. In addition, when conditioning on eligibility, spillover effects and direct/total effects may not be very comparable. In future studies, we recommend planning to measure within-cluster spillover measurement during the study design phase so that measures to minimize bias and maximize study validity can be taken (e.g., by matching individuals in the control group to comparable individuals not receiving treatment in the treatment group).

Distance-based spillovers

Four papers estimated distance-based spillover parameters, and each measured a slightly different parameter (Table 8, parameters 13-17) (Hawley et al. 2003; Björkman and Svensson 2009; Joshi and Schultz 2013; Banerjee et al. 2013). In our overall risk of bias assessment, two of these studies had high quality evidence, one had moderate quality evidence, and one had very low quality evidence. Several included papers estimated spillovers conditional on treatment density that were indexed by distance. In this section, we focus on papers that only assessed the effect of distance – not treatment density – on spillovers.

Rigorous designs for estimating distance-based spillovers

Two studies in the review utilized rigorous approaches to estimate distance-based spillovers that we recommend for use in future studies. Banerjee et al. utilized a cluster-randomized design to measure spillovers between clusters. As discussed in Section 3.2.2, investigators enrolled villages within 6 km of villages that were randomized to receive an immunization

campaign with incentives (Table 8, parameter 13) (Banerjee et al. 2010). They found evidence of spillovers; the relative risk for the number of immunizations in villages neighboring those that received an immunization campaign was 1.18 (95% CI 0.92, 1.43), and it was 1.48 (95% CI 1.18, 1.77) in villages neighboring those that received an immunization campaign with incentives.

Another study design that generated high quality evidence of distance-based spillovers measured outcomes within different distances from villages that were randomized to receive insecticide-treated nets or to control (Table 8, parameters 14-15) (Hawley et al. 2003). Investigators assessed possible spillover effects by comparing malaria, anemia, and child mortality in households within 0-299m, 300-599m, 600-899m, and ≥900m of the nearest intervention and control villages. They found that malaria, anemia, and child mortality were lower in compounds without ITNs who lived <300m from villages receiving ITNs compared to those who lived ≥900m from such villages.

Both of these studies require an assumption that there were no negative spillovers from the control areas to intervention areas. Without this assumption inference is greatly complicated. An alternative approach, to be discussed further below, utilizes statistical models to estimate spillovers when it is possible that spillovers occurred in the control group (i.e., contamination occurred).

Other designs for estimating distance-based spillovers

There were two other approaches used in included studies that measured distance-based spillovers. These designs produce lower quality evidence, so we only describe them briefly. One study estimated the total effect conditional on whether treatment and control areas were within a specific distance of each other (Table 8, parameter 16) (Björkman and Svensson 2009). While this approach can produce internally valid results if areas were randomized, it provides less rigorous evidence than the approaches described above because it does not isolate the spillover effect from direct effects and does not assess a potential dose-response pattern over greater distances. Another study leveraging a previous cluster-randomized trial assessed whether spillovers occurred into the control group by comparing outcomes among individuals on the boundaries of the comparison areas closest to the program areas (Table 8, parameter 17) (Joshi and Schultz 2013). Ideally, even when spillovers are of interest, intervention and control units will have sufficient physical or social distance that they can be considered independent in order to allow for accurate spillover estimation. Thus, we do not recommend this approach in future studies.

Association between distance and spillovers

We aimed to assess whether spillover magnitude was associated with the distance over which they were measured. Only two of the four papers focusing on distance-based spillovers reported the typical distance between independent units. Between these two papers, the spillovers estimated in the study that assessed them on a smaller scale (6km) were positive and

statistically significant (Banerjee et al. 2010), whereas those estimated in a study that assessed them on a larger scale (30km) were not (Björkman and Svensson 2009).

Measuring spillovers into the control group

Two studies (Björkman and Svensson 2009; Joshi and Schultz 2013) estimated spillovers into the control group. In studies that hypothesize that spillovers may occur in the control group, the control group is "contaminated" and can no longer be used as a valid counterfactual. In this case, estimation is inherently more complicated because treated and control units can no longer be assumed to be independent. As a result, one cannot rely on randomization to derive inference, and a statistical model is needed to estimate spillovers. The approach employed by Banerjee et al. is preferable (Banerjee et al. 2010); by prospectively building in spillover measurement, they were able to reasonably assume that no contamination was present in their control clusters, and they could rely on randomization-based, non-parametric inference to estimate the spillover effects.

In summary, a variety of different designs producing evidence of varying quality were used in included studies to estimate distance-based spillovers. We recommend that future studies that estimate this parameter consider the designs used by Banerjee et al. (Banerjee et al. 2010) and Hawley et al. (Hawley et al. 2003), which can allow for rigorous estimation of spillover effects. While the studies discussed in this section focused on physical distance, such approaches could also be implemented in studies focusing on spillovers over social distances as well. We recommend that authors pre-specify how they plan to measure distance, report their rationale for their choice of distance measures, and assess spillovers over multiple distances if possible in order to assess whether spillovers decay over increasing distances.

Spillovers conditional on density

Seven studies estimated spillover effects conditional on treatment density (Table 8, parameters 18-22) (Miguel and Kremer 2004; Dupas 2006; Ziegelhöfer 2012; Bhattacharya et al. 2013; Baird et al. 2013b; Chong et al. 2013; Perez-Heydrich et al. 2014), and one study estimated effects conditional on outcome density (Godlonton and Thornton 2012). Table 9 lists features of these studies within different categories of parameters. At first glance these parameters may appear similar to the treatment coverage mean/effect. However, they are distinct in that they typically do not compare effects across numerous levels of treatment coverage, and they typically adjust for treatment coverage using interaction terms with treatment indicators in a statistical model. These studies estimated spillovers of a wide range of interventions, including those occurring through physical proximity (e.g., school-based deworming) and those occurring through social channels (e.g., sexual health education). One of these studies produced high quality evidence, five had moderate quality evidence, and two had low quality evidence.

Definition of treatment density

In studies estimating this parameter, treatment density was frequently measured by estimating the proportion receiving treatment within a specific physical distance around households or clusters (Miguel and Kremer 2004; Ziegelhöfer 2012; Bhattacharya et al. 2013; Baird et al. 2013b). To assess spillovers over social distance, one study estimated density using the

proportion of close friends of study participants who participated in an intervention (Chong et al. 2013). Another study used a spatial clustering method to define neighborhoods around study households (Perez-Heydrich et al. 2014). All but one study estimated treatment density in populations where treatment was originally randomized. This approach minimizes bias by ensuring that the distribution of treated units is not associated with the outcome.

Rigorous study designs to estimating spillovers conditional on density

One study utilized a causal inference approach to estimate spillovers conditional on density in a population that was previously randomized to receive a cholera vaccine (Perez-Heydrich et al. 2014). The original study was randomized, so to account for the presence of spillovers in the study, investigators weighted individuals' potential outcomes by the inverse of the probability of participating in the trial. They then compared the average inverse-weighted risk of cholera at differing levels of vaccination coverage. This approach can produce high quality evidence by accounting for potential systematic differences between individuals who choose to receive the vaccine or not. Furthermore, the parameter utilizes outcome data at the individual level, allowing for stronger scientific inference than the studies using the same dataset that estimated ecologic parameters using the treatment coverage mean or treatment coverage effect (Ali et al. 2005; Emch et al. 2006; Ali et al. 2008; Emch et al. 2009; Root et al. 2011).

Studies building on the design used by the Miguel and Kremer deworming study

The approach used by Miguel and Kremer (Miguel and Kremer 2004) to estimate spillovers between clusters has been replicated and modified in numerous other studies, including some in this review (Ziegelhöfer 2012; Baird et al. 2013b). Miguel and Kremer defined a statistical model that included indicators for school treatment assignment, individuals' eligibility for treatment, the number of children receiving treatment in schools assigned to receive treatment within 0-3 km and 3-6 km, and the total number of children within 0-3 km and 3-6 km (Table 8, parameter 18b). By conditioning on the number receiving treatment and the total number of individuals within specific distances, the parameter effectively conditions on the treatment density surrounding each study child. For the average number of pupils in their study population, they found a 12% reduction in moderate to heavy worm infections for children 0-3 km away from schools receiving the deworming program and an 11% reduction for children 3-6 km away from such schools. In the original analysis, both these findings were statistically significant.

As discussed in section 5.3.1.2, a replication of this study was conducted, and it identified errors in the coding of the population density variable. Once this error was corrected, the spillover for children 0-3 km away was 9% (0.21 x 448/1000) and the spillover for children 3-6 km away was 6% (0.05 x 1108/1000). Following this change, the spillover for children 0-3 km away was statistically significant, but the 3-6 km spillover was not (Aiken et al. 2015). In their response to the replication study, Hicks, Miguel, and Kremer explain why the recoding of this variable resulted in poor precision of spillover estimates between schools 4-6 km away and emphasize the policy relevance of the finding of spillovers within 0-3 km (Hicks et al. 2014). We refer readers to the thorough discussion of the replication study's findings for more details (Davey et al. 2014; Hicks et al. 2014; Aiken et al. 2015).

Many studies estimating spillovers conditional on treatment density adapted Miguel and Kremer's statistical model (Godlonton and Thornton 2012; Ziegelhöfer 2012). Using statistical models to control for treatment density is subject to several limitations. One is the assumption of linearity; if a single continuous measure of treatment density is conditioned on, most standard parametric approaches will assume a linear association between the outcome and treatment density. By including indicators for the proportion of individuals that received treatment within 0-3 km and 3-6 km, Miguel and Kremer allowed for different effect sizes over those two distance ranges, but within those ranges, the treatment effect assumed a linear model. Whether or not such assumptions are reasonable for other interventions and outcomes depends on the hypothesized scale and mechanism of spillovers. Another consideration in modeling spillovers conditional on treatment density is whether the model requires extrapolation beyond the support in the data. For example, if a model conditions on treatment density but no cluster in the study had 0% treatment density, using model parameters to estimate the effect under 0% treatment density requires extrapolation using a parametric model, which could create bias if the model is mis-specified. Defining parameters that can be identified with the observed data helps to avoid this type of model extrapolation. For example, one could examine the empirical distribution of treatment density and then define a parameter with a counterfactual at an observed level of treatment density that allows for a meaningful contrast.

In summary, only one study measuring spillovers conditional on density measures produced high quality evidence. However, future studies can estimate spillover density parameters rigorously if the area in which density is measured is carefully defined and pre-specified. Randomizing treatment is preferable because it ensures that the distribution of treatment is independent of the outcome or other confounders. These parameters can aid program implementers in understanding how to distribute programs in order to yield the greatest health impact and highest cost-effectiveness.

Spillovers conditional on exposure to cases

Two papers estimated spillover parameters conditional on exposure to cases, both of which evaluated vaccine efficacy (Table 8, parameters 23-25) (Préziosi and Halloran 2003; Baptista et al. 2006), and a third paper produced data which could be used to estimate vaccine efficacy but did not estimate the parameter (Paul et al. 1962). These studies enrolled pertussis cases through surveillance systems and then enrolled household members of cases with no history of pertussis as controls. Préziosi and Halloran estimated VE_S=34%, and Baptista et al. estimated VE_S=12.5% (95%CI -5.3, 27.3). These results indicate that pertussis vaccination did not exert a significant protective effective for unvaccinated household members. VE_I was estimated to be 85% (95%CI 46, 95%) by Préziosi and Halloran and 61.6% (95%CI 12.8, 83.1%) by Baptista et al., indicating strong protection against transmission of pertussis from vaccinated cases to controls. In summary, these parameters were only estimated in studies evaluating vaccines. While they allow for rigorous estimation of spillover effects, they are most appropriate for estimation using case-control designs. This approach is well suited to health outcomes measured through surveillance systems in a population in which most individuals were offered the intervention. When the intervention is targeted to a specific sub-population and outcome measurement is prospective, the case-control design and these vaccine efficacy parameters are less appropriate.

Social network spillovers

Three papers measured spillovers through social networks, each estimating different parameters; all produced moderate quality evidence (Kremer and Miguel 2007; Tontarawongsa et al. 2011; German et al. 2012).

Spillovers among social network members

A rigorous design used by German et al. randomized an intervention to individuals and then enrolled peers of individuals assigned to intervention and control (Table 8, parameter 26) (German et al. 2012). They did not find evidence of spillovers of a peer network intervention among social network members.

Spillovers conditional on knowing individuals who received an intervention

Two studies estimated social network spillovers by conditioning on whether respondents knew individuals who received treatment. Miguel and Kremer assessed whether children whose parents had more social links to schools randomized to receive a deworming program were more likely to take deworming (Table 8, parameter 27) (Kremer and Miguel 2007). Contrary to expectation, they found that each additional social link with a school allocated to treatment was associated with a 3.1% (95%CI 0.4%, 6%) decrease in the probability that a child took deworming. Tontarawongsa et al. found that the fraction of household members sleeping under an ITN was 7% higher (SE=0.042) among households that knew all participants in a cluster-randomized trial that received free ITNs compared to those who knew no participants (Tontarawongsa et al. 2011). They did not find spillovers for net or ITN acquisition. They also assessed whether social network members' average per capita bed nets was associated with ITN purchases and ITN use and did not find evidence of spillovers.

Overall, there was little evidence to support spillovers through social networks for a school deworming program and peer network intervention, but there was some evidence that providing free insecticide-treated nets led to spillovers among social network members. In each of these studies outcomes were self-reported, participants or outcome assessors were not blinded, and we were uncertain about allocation concealment. Thus, we found no rigorous evidence of social network spillovers in these three studies.

An important aspect of social network spillovers that affects their policy and program relevance is the contrast built into the parameter definition. One could estimate spillovers associated with each additional social contact, or one could estimate spillovers associated with knowing everyone in a network compared to knowing no one. For example, Tontarawongsa et al. estimated the latter contrast in their ITN study (Tontarawongsa et al. 2011). Parameters that estimate spillovers with more realistic contrasts are likely to be more program- and policy-relevant.

5.3.2 Results within intervention categories

Vaccines

Twenty-two included studies estimated spillover effects of vaccines for cholera (n=8) (Ali et al. 2005; Emch et al. 2006; Ali et al. 2008; Emch et al. 2009; Root et al. 2011; Khatib et al. 2012;

Ali et al. 2013; Perez-Heydrich et al. 2014), pneumococcal conjugate (n=5) (Roca et al. 2011; Egere et al. 2012; Roca et al. 2013; Hammitt et al. 2014; Root et al. 2014), pertussis (n=3) (Cooper and Fitch 1983; Préziosi and Halloran 2003; Baptista et al. 2006), *Haemophilus influenzae* type b (n=2) (Forleo-Neto et al. 1999; Chen et al. 2014), typhoid (n=2) (Sur et al. 2009; Khan et al. 2012), polio (n=1) (Paul et al. 1962), and BCG, DPT, polio and measles (n=1) (Banerjee et al. 2010). In all these studies, the most plausible mechanism of spillover is physical proximity due to the nature of transmission for vaccine-preventable diseases. Of these studies, 4 had high quality evidence, 13 had moderate quality evidence, and 5 had low or very low quality evidence.

Eleven studies evaluated spillovers of vaccines by comparing incidence rates among subpopulations with differing levels of vaccine coverage (Cooper and Fitch 1983; Forleo-Neto et al. 1999; Ali et al. 2005; Emch et al. 2006; Ali et al. 2008; Emch et al. 2009; Root et al. 2011; Khatib et al. 2012; Ali et al. 2013; Chen et al. 2014; Root et al. 2014). Ten of these studies, which evaluated cholera, pertussis, *Haemophilus influenzae* type b, and pneumococcal conjugate vaccines, reported a decrease in incidence associated with increased vaccine coverage among unvaccinated individuals and/or all individuals in the study population (Cooper and Fitch 1983; Forleo-Neto et al. 1999; Ali et al. 2005; Emch et al. 2006; Ali et al. 2008; Emch et al. 2009; Root et al. 2011; Khatib et al. 2012; Ali et al. 2013; Chen et al. 2014).

Spillover effects were larger in studies that measured spillovers on smaller scales. For example, protective efficacy of the typhoid vaccine in clusters of approximately 700 people was 44-45% (Sur et al. 2009), whereas the protective efficacy of pertussis vaccine within households was 61.6% (Baptista et al. 2006) and 85% (Préziosi and Halloran 2003). No negative, statistically significant spillovers were reported in these studies. In addition, spillover effects were stronger among studies with higher vaccine coverage. For example, Egere et al. reported a 70% reduction in vaccine-type pneumococcus in villages with complete pneumococcal conjugate vaccine coverage compared to villages in which only young children were vaccinated (Egere et al. 2012).

Overall, the majority of studies evaluating spillovers of vaccines included in this review found evidence of reduced disease among nonrecipients. Most of these studies had moderate or high quality. Evidence of spillovers was less common among studies that estimated spillovers with parameters less likely to be biased due to confounding (e.g., the within-cluster spillover effect). The finding that spillovers were larger in studies assessing spillovers on smaller scales and when vaccine coverage was higher is consistent with the theory of disease transmission for vaccine-preventable diseases. Since transmission occurs through physical proximity, we would expect larger spillovers on smaller scales and when coverage is higher. An important caveat is that there is likely substantial dependence among vaccine study findings – 6 out of the 18 included vaccine studies re-analyzed data from the same cholera vaccine trial in Matlab, Bangladesh. In addition, three studies evaluated the pneumococcal conjugate vaccine in The Gambia, and two studied the same population, so these studies' findings are likely to be dependent.

Mass drug administration for parasite control

Mass drug administration interventions aim to eliminate parasite infection of uninfected individuals by treating large populations in order to interrupt disease transmission. Spillovers may occur among individuals in populations targeted for mass drug administration who were not eligible for treatment or who did not receive treatment due to incomplete coverage. Seven studies evaluated mass drug administration interventions for parasite control (Miguel and Kremer 2004; Chidambaram et al. 2004; Kremer and Miguel 2007; House et al. 2009; Baird et al. 2013b; Ozier 2014; Shekhawat et al. 2014). One of these studies produced high quality evidence overall, three had moderate quality, and four had moderate or low quality evidence.

These include mass administration of azithromycin to control trachoma (Chidambaram et al. 2004; House et al. 2009; Haile et al. 2013; Shekhawat et al. 2014) and mass administration of benzimidazoles (i.e., deworming) to control soil-transmitted helminth infections (Miguel and Kremer 2004; Baird et al. 2013b; Ozier 2014). One study evaluated exposure to information about deworming in the context of a school-based deworming program (Kremer and Miguel 2007). The primary spillover mechanism for these interventions is physical proximity. Soil-transmitted helminth infections are transmitted when an infected person passes helminth eggs through their stool and an uninfected person is exposed to the eggs due to fecal contamination of the environment. Thus, transmission is more likely to occur when an individual lives in close proximity with other infected individuals. Trachoma is spread through interpersonal contact, shared clothing, and flies that come into contact with the eyes or nose—thus physical proximity is also a major driver of transmission.

Three of the four studies of mass azithromycin to control trachoma for which quantitative spillover estimates were available found positive, statistically significant spillovers. In Ethiopia, mass azithromycin treatment resulted in a 35% (95%CI 8%, 55%) decrease in trachoma among individuals who did not receive treatment in clusters randomly allocated to treatment 12 months after mass treatment (House et al. 2009). These confidence intervals present a best-case scenario since they are not adjusted for clustering. A similar study in Ethiopia found a 2.9-fold (95% CI 1.1, 7.5) reduction in the odds of trachoma among those not receiving azithromycin who lived in areas where mass azithromycin treatment occurred (Chidambaram et al. 2004). All studies evaluating control of trachoma through mass drug administration were conducted in Ethiopia, so there is possible dependence between these findings.

Of the four studies evaluating mass deworming to control soil-transmitted helminth infections, three (Kremer and Miguel 2007; Baird et al. 2013b; Ozier 2014) were based on the study of school-based deworming conducted by Miguel and Kremer (Miguel and Kremer 2004). Miguel and Kremer reported spillovers of school-based deworming to untreated students in treated schools as well as spillovers among pupils of schools within 0-3 km (Miguel and Kremer 2004). The other three studies evaluating school-based deworming assessed spillover effects for other outcomes (self-reported health, body mass index, child growth, deworming uptake) and thus are not directly comparable.

In summary, there is evidence of spillovers of mass administration with azithromycin for trachoma control in Ethiopia, but evidence from studies conducted in other countries would strengthen the generalizability of these findings. Evidence of spillovers of mass deworming administration on helminth infection is restricted to a single study.

Health education

Five studies evaluated educational programs focusing on reducing neonatal mortality (Azad et al. 2010), reducing sexually transmitted infections (Dupas 2006; Chong et al. 2013), and improving child nutrition and growth (Singh 2011; Fitzsimons et al. 2012). None of these studies produced high quality evidence, four produced moderate quality evidence, and one produced low quality evidence. These studies evaluated spillovers that occurred through a variety of mechanisms including physical proximity, social proximity, learning and imitation, and norm shaping.

An evaluation assessed spillovers of a program to promote women's groups in Bangladesh as a means of providing education about safe delivery and neonatal health (Azad et al. 2010). Neonatal mortality was 7% lower among women in villages with women's groups who did not participate compared to that among women in control villages without the women's groups, suggesting possible spillover effects. The authors assessed numerous other outcomes as well, and there was not evidence of spillovers for all outcomes. The authors did not discuss potential mechanisms of spillover; plausible mechanisms include social proximity, learning and imitation, and/or norm shaping.

Fitzsimons et al. conducted an evaluation of a program providing information on infant health and child nutrition to mothers in Malawi to determine whether older children in the same households who were not targeted by the program benefited (Fitzsimons et al. 2012). They did not find any statistically significant spillovers for child growth indicators, diarrhea, vomiting, or other symptoms among these older children. Singh evaluated a nutrition information program targeted at mothers and assessed whether spillovers occurred for children of mothers who did not participate in the program but who lived in program areas (Singh 2011). There was no evidence of spillovers on child growth indicators.

Two studies measured spillovers of school-based health education programs that aimed to increase knowledge of sexually transmitted infections and reduce risky sexual behavior. One study found that sexual health knowledge scores were lower for students who had no friends in an online sexual health education program compared to scores for students for whom all their friends participated in the course (Chong et al. 2013). There was no evidence of spillover effects among students who did not participate in the course but who attended schools where the course took place. Another study evaluated whether a program providing information about HIV transmission resulted in less risky sexual behavior (Dupas 2006). Schools were randomly selected to receive the program. To assess spillovers, investigators compared health behaviors among students in schools with different proportions of students participating in the program. Investigators found that the proportion of students allocated to treatment in a school was associated with condom use for girls but not boys. For both boys and girls findings were not statistically significant when evaluating whether the student had sex without a condom.

In summary, evidence of spillovers of health education programs is not consistent across studies and settings, and the quality of evidence is moderate at best. There was some evidence of spillovers of such programs within a village and school, but not within households. Unlike spillovers resulting from physical proximity, spillovers of health education programs resulting

from social mechanisms appear not to be associated with the physical scale on which spillovers are measured.

Cash transfers

Five studies evaluated conditional cash transfer (CCT) and unconditional cash transfer (UCT) programs in Colombia, Malawi, Mexico, and Paraguay (Handa et al. 2001; Ribas et al. 2011; Avitabile 2012; Baird et al. 2013a; Contreras and Maitra 2013). We classified the quality of evidence from these studies as moderate for two studies and low or very low for three studies. In these studies, the purported spillover mechanisms were primarily learning, imitation, norm shaping, and social proximity.

A few studies assessed cash transfer programs' impact on preventative health behaviors. For conditional cash transfer programs, there was evidence of increased health screening (e.g. nutrition surveillance, cancer screening) among ineligible individuals in two studies (Handa et al. 2001; Avitabile 2012); one study found no such spillovers for screening (Ribas et al. 2011). A possible mechanism for these findings is that increased health screening and improved health behaviors resulting from the conditional cash transfer program may have altered social norms, or non-participants may have learned these health behaviors from participants.

Two studies assessed cash transfer programs' effect on health. One study found evidence that ineligible individuals in areas where others received conditional cash transfers experienced less illness, but these effects were not sustained after three years for all ineligible individuals (Contreras and Maitra 2013). One study evaluated whether a conditional and unconditional cash transfer program that targeted adolescent girls reduced psychological distress (Baird et al. 2013a). While they found 8-14% reductions in psychological distress among participants during the program, there was evidence that nonparticipants experienced a 6.4% increase in psychological distress (i.e., a negative spillover). To explain these negative spillovers, the authors hypothesized that adolescent girls' utility is a function not only of their own income but also their relative income.

Among studies included in this review, cash transfers were associated with increased preventive health screenings and some increases in health among non-recipients but the quality of evidence from these studies was low. Two of these studies were based on the same conditional cash transfer program (PROGRESA), so these studies' findings are likely to be dependent.

HIV/AIDS counseling and treatment

One study evaluated an HIV/AIDS voluntary counseling and testing program. They hypothesized that spillovers may have occurred through social proximity; specifically, they assessed whether the proportion of nearby neighbors who received HIV test results was associated with choosing to learn one's HIV test results (Godlonton and Thornton 2012). They found positive spillovers: a 10% increase in neighbors who learned their HIV results was associated with a 1.1% increase in the probability that an individual sought their own HIV test results.

Another study evaluated whether a parent's HIV/AIDS treatment affected their child's nutritional status (Zivin et al. 2009). They hypothesized that improved health and productivity resulting from HIV/AIDS treatment could allow for increased spending on child nutrition via income and substitution effects. In addition, they hypothesized that such spillovers were more likely to occur among parents who were on HIV treatment for more than 100 days compared to those on treatment for less than 100 days due to the time it takes for the treatment to improve parent health and productivity. They did not find strong evidence of such spillovers when comparing weight-for-height z-scores of children whose parents had been on HIV/AIDS treatment for >100 days vs. <100 days.

Because only two studies evaluated interventions related to HIV testing and treatment and both were of moderate quality, we cannot make conclusions about the evidence of spillovers for this category of interventions.

Insecticide-treated nets

Three studies evaluated programs that distributed insecticide-treated nets (ITNs) for free in Kenya (Hawley et al. 2003), with subsidies in Kenya (Bhattacharya et al. 2013), and with education and microloans in India (Tontarawongsa et al. 2011). One study had high quality evidence, one had moderate quality evidence, and one had low quality evidence. One study found evidence of notable reductions in malaria and anemia among individuals who did not receive ITNs and lived within 300 m of villages that received free ITNs compared to those who lived ≥900 m from such villages; however, spillovers were not present for child mortality (Hawley et al. 2003). Another study found that the probability of an ITN purchase was associated with the number of people eligible to receive subsidies in nearby areas (Bhattacharya et al. 2013). In an assessment of a program offering ITNs with education and microloans, investigators found that the fraction of household members having slept under an ITN the previous night was associated with the number of people participating in the program, but the program did not result in spillovers increased net or ITN acquisition among nonparticipants (Tontarawongsa et al. 2011).

In summary, the few studies that have estimated spillovers of ITN programs with and without subsidies suggest that they lead to positive spillovers of health outcomes for individuals in nearby areas who did not receive free ITNs. The association between spillovers and distance to treated areas suggests that physical proximity is a major mechanism of spillovers of ITNs. It is also possible that social proximity resulted in acquisition of ITNs by individuals nearby program areas.

Nutrition

Two studies estimated whether siblings of children that received free meals at school and take home meals in Burkina Faso and Laos experienced spillovers (Kazianga et al. 2009; Buttenheim et al. 2011). One study had moderate quality evidence, and one had low quality evidence. The mechanism of spillover hypothesized in both studies was substitution: if a fixed amount of resources are available to purchase food for a household, provision of free meals at schools may free up resources to feed siblings more food at home. One study found that weight-for-age was higher among pre-school aged siblings of children who received take home

rations from school, but that there were no spillovers for siblings of children who received free meals at school with no take home meals (Kazianga et al. 2009). The second study mentioned that they found some evidence of spillovers among younger and older siblings of children participating in a school feeding and take home ration program, but they did not present disaggregated results for spillovers (Buttenheim et al. 2011). In these two included studies, there was weak evidence of spillovers on siblings of children participating in school feeding programs.

Maternal and child health

Two studies evaluated spillovers in a maternal and child health program in Matlab, Bangladesh which provided family planning services, basic health education, antenatal care, and safe delivery kits (Chaudhuri 2005; Joshi and Schultz 2013). Both of these studies produced very low quality evidence. One study estimated spillovers on body mass index of elderly women and adult men who were not targeted as part of the program. They hypothesized that income effects may occur in which resources provided by the program free up household resources to be spent on food for other individuals. They also hypothesized that health and hygiene information provided to mothers may also benefit other individuals in the household. They found spillovers for elderly women's body mass index but not for adult or elderly men.

Another study evaluating the same program estimated spillovers by comparing outcomes of individuals in the boundaries of control areas near treatment areas to those of individuals in control areas further from treatment areas. They found that women in boundary control areas had 0.35 fewer children on average. There was some weak evidence of higher child mortality and poorer self-reported health in boundary areas.

Both of these studies evaluated the same program and utilized similar datasets, so their results are likely to be dependent. Their results should be interpreted with caution since both studies had very low quality evidence.

Water and sanitation

Three studies assessed the effect of water sources, water filtration, and improved sanitation on health. Two of these studies produced moderate quality evidence, and one produced very low quality evidence. One study assessed whether the fraction of people whose community received water points was associated with diarrhea (Ziegelhöfer 2012). They did not find statistically significant evidence of such spillovers. Another study found that communities with a higher proportion of sari or other cloth filtration of water had a lower incidence of cholera (Huq et al. 2010). Finally, a study assessed spillovers in a mass azithromycin distribution and latrine construction program in Ethiopia. They found that a 10% increase in latrine use in study areas was associated with a 2% decrease in trachoma infection (95% CI 0.2, 3.9) (Haile et al. 2013). The latter two studies' comparison of health at different intervention coverage levels jointly measures both direct and spillover effects and indicates that spillover effects may be present.

Other interventions

Other interventions for which spillovers were measured include a women's empowerment program (Janssens 2005), a peer support intervention for drug users (German et al. 2012), and community monitoring of health services (Björkman and Svensson 2009).

Janssens et al. evaluated a women's empowerment program in India and assessed whether non-participants in villages where the program was conducted experienced any positive spillovers (Janssens 2005). They found that non-participants in program villages were approximately 12-27% more likely to vaccinate their children than those in control villages. Spillover effects were nearly as large as the effects among program participants; for immunization of children, the spillover effects were 40-54% of the program effect.

As discussed above, German et al. evaluated a peer support intervention among drug users to assess whether their peers who were not participating in the program experienced decreases in depression. Peers of intervention participants experienced a 9.5% decrease in depression as measured by the Center for Epidemiologic Studies Depression Scale, and those connected to control participants experienced a 9.2% decrease. However, these results were not statistically significant at the alpha=0.05 level.

Björkman and Svennson evaluated whether a program encouraging community monitoring of health services in Uganda resulted in spillovers into the control group; clinics were assigned to treatment or control (Björkman and Svensson 2009). They estimated the program effect conditional upon an indicator for whether a control clinic was within 10 km of the nearest treatment clinic and did not find any evidence of spillovers into the control group using this approach.

5.4. Publication Bias

We hypothesized that publication bias might be more likely for spillover effects because they were not pre-specified in most studies, so investigators may have been more likely to mention spillovers in publications if they found they were present. Figures 15a-d present funnel plots of spillover and total/direct effects. We produced separate funnel plots for risk differences vs. risk ratios because the standard errors for each could not be converted to a single scale with the information available in most studies.

In the funnel plot for spillover effects estimated with risk differences (Figure 15a) the points are well-balanced around 0, indicating minimal publication bias, whereas in the plot for spillover effects estimated with risk ratios (Figure 15b) the majority of points are for risk ratios less than 1 ("positive spillovers"), indicating strong publication bias. These patterns are confounded by academic discipline; the risk ratio plot only contained estimates from public health papers, most of which estimated spillovers of vaccines or mass drug administration programs. While the risk ratio spillover plot suggests that positive spillovers are more likely to be reported, the interventions included in that plot – vaccines and MDA – are unlikely to result in negative spillovers. On the other hand, the risk difference plot mostly contains estimates from economics papers evaluating a range of different interventions for which the direction of spillover effects to be expected is less clear. In addition, economics papers tended to report results for a larger number of different model specifications and subgroups than public health papers, which may contribute to the wider range of positive and negative spillovers reported. Figures 15c and 15d

present funnel plots for total and direct effects, and in both of them points are evenly distributed across the null value suggesting little publication bias. On the whole, the trends in Figures 15a-d are consistent with our expectations: publication bias appears to be more common for spillover effects than total or direct effects.

Of 495 separate spillover parameters estimated in 42 studies that estimated statistical significance, 36% were statistically significant at the 95% confidence level. For 110 analyses that were pre-specified, 39% were statistically significant versus 61% (n=374) of analyses that had not documented whether they were pre-specified. The finding that studies that did not pre-specify spillover estimation were twice as likely to have statistically significant findings also suggests that publication bias was likely present in the studies included in this review.

5.5. Additional Analyses

We identified 14 terms commonly used to describe the concept of spillovers (Table 10). The most common terms were “indirect effect” and “spillover” followed by “externality/externalities”. “Indirect protection” and “herd protection” were other common terms.

6. Discussion

6.1. Summary of Main Results

In this review, we have summarized methods for estimating spillover parameters across disciplines and synthesized them into a unified framework. To our knowledge, this is the first systematic review of health-related spillovers and interdisciplinary review of spillover estimation methods. The Inter-American Development Bank published a report that provides an overview of spillover effects within program evaluations and brief guidelines for estimating one class of spillover parameters (Angelucci and Maro 2010). The report did not include a comprehensive review of primary research or possible design choices nor did it formally identify the assumptions needed for causal inference in the presence of spillovers. Baird recently published a working paper focused on estimating spillovers that included identification assumptions, however, the paper only focuses on the double-randomized design and its extensions (Baird et al. 2014). Halloran and others have made extensive methodological contributions to the estimation of indirect effects and transmission in the context of vaccine studies, particularly with respect to causal inference and identification assumptions, however, their studies are focused only on vaccines (Halloran and Struchiner 1991; Longini et al. 1998; Hudgens and Halloran 2008; Halloran et al. 2010; VanderWeele et al. 2012; Tchetgen and VanderWeele 2012). To date, discussion of identification of spillover parameters has largely occurred within academic disciplines with minimal cross-referencing between disciplines. Here, we synthesized results across three academic disciplines and defined spillover parameters within a unified framework.

6.1.1 Common spillover parameters

The most frequently measured parameter was within-cluster spillover effects and accompanying total and direct effects. This parameter is likely common because it is easily estimated using a cluster-randomized design. Only two studies utilized a double-randomized design, which is the ideal design for estimation of within-cluster spillovers (Baird et al. 2013a; Chong et al. 2013). A common limitation of studies measuring within-cluster spillovers was a lack of clear description of buffer zones between treated and control clusters. Such information is needed to assess whether the control group can serve as a valid counterfactual.

Another frequently measured parameter was the treatment coverage mean or treatment coverage effect. As discussed above, these parameters estimate ecological associations and thus are inherently a less rigorous tool for estimating spillovers than alternative parameters. In addition, many of the studies estimating treatment coverage parameters did not adjust for potential confounders or present measures of precision. These parameters are relatively easy to estimate with existing data and are an appropriate tool for initial assessments of possible spillovers that can later be assessed with more rigorous methodology.

Nearly all of the studies in the public health literature assessed spillovers via geographic proximity, whereas social mechanisms were more commonly explored in the economics literature. Only three studies measured social network parameters (Kremer and Miguel 2007; Tontarawongsa et al. 2011; German et al. 2012). As mentioned above, one limitation of measuring social network spillovers is that doing so can require a near census of individuals,

which is time consuming to collect and not always feasible. Assessing spillover through geographic proximity may be an appropriate proxy in some populations. Future studies could implement validity assessments to assess how well geographical distance is associated with social distance as Giebultowicz et al. have done in studies of cholera in Bangladesh (Giebultowicz et al. 2011).

6.1.2 Summary of spillover evidence for different interventions

Here, we briefly summarize findings by intervention category. We interpret spillover findings within specific categories of interventions with caution because with the exception of vaccines, there were very few studies measuring spillovers of any given intervention. Furthermore, for each intervention, many of the studies produced moderate or low quality evidence, so spillover findings must be interpreted with great caution.

- **Vaccines**: Most of the studies evaluating spillovers of vaccines found evidence of reduced disease among unvaccinated individuals who lived with or near to other vaccinated individuals. There was evidence that spillovers were larger in studies assessing spillovers on smaller scales (e.g., spillovers were stronger in households than in villages) and when vaccine coverage was higher. An important caveat is that there is likely significant dependence among vaccine study findings – 8 out of the 22 included vaccine studies re-analyzed data from the same two study populations.
- **Mass drug administration for parasite control**: There is evidence of spillovers from three studies of mass administration of azithromycin for trachoma control. However, 2 of 3 studies of this intervention were conducted in Ethiopia. Evidence from studies in other populations would strengthen the generalizability of these findings. There was also evidence of within-school and between-school spillovers in a study of school-based deworming in Kenya.
- **Health education**: Evidence of spillovers of health education programs is not consistent across studies and settings, and the quality of evidence is moderate at best.
- **Conditional and unconditional cash transfers**: There was only low quality evidence assessing spillovers of cash transfer programs.
- **HIV counseling, testing, and treatment**: Only two studies, both of moderate quality, evaluated interventions related to HIV testing and treatment.
- **Insecticide-treated nets**: There is evidence from one high quality study that ITN programs lead to reduced disease among individuals in nearby areas who did not receive free ITNs. There is only mixed, moderate to low quality evidence about whether ITN programs offering vouchers or cash transfers result in spillovers on ITN acquisition.
- **School feeding programs**: There was weak evidence of spillovers from two studies of low and moderate quality among siblings of children participating in school feeding programs.
- **Maternal and child health**: Only two studies assessed spillovers of maternal and child health programs, and both had very low quality evidence.

- **Water and sanitation**: A study producing moderate quality evidence found no association between the fraction of people whose community received water points and diarrhea. Another study of moderate quality found that communities with a higher proportion of sari or other cloth filtration of water had a lower incidence of cholera.

6.2. Overall Completeness and Applicability of Evidence

We summarized the existing literature measuring spillover effects of interventions in low- and middle-income countries on health outcomes. Fifty-four studies conducted in 20 low- and middle-income countries met our inclusion criteria. We identified 29 different spillover parameters estimated in the included studies using a wide range of randomized and observational designs. In our assessment of the quality of each study's evidence, we classified 6 as high quality, 30 as moderate quality, 12 as low quality, and 6 as very low quality.

6.3. Quality of the Evidence

6.3.1 Main risks of bias

There are two overarching sources of bias in the studies in this review. The first is related to the allocation of treatment. For any spillover parameter, ideally treatment would be double-randomized. Utilizing two-stage randomization minimizes selection bias and unmeasured confounding. Only 2 out of the 54 studies utilized double-randomized designs (Baird et al. 2013a; Chong et al. 2013). In some cases, a design can randomize clusters without randomizing individuals within treatment clusters and still make valid inferences. For example, the cluster-randomized design used by Banerjee et al. in which villages near to treated villages were enrolled minimizes bias as well as violations of SUTVA (Banerjee et al. 2010). However, in the majority of studies, individuals enrolled to measure spillovers were not randomly assigned and may be systematically different from those directly receiving an intervention.

Second, many studies did not pre-specify spillover measurement. When spillover estimation is not pre-specified, publication bias is more likely. In addition, spillover parameters may be defined in a way that increases the chance of detecting positive spillovers, whether intentionally or not. Many studies estimated spillovers conditional on specific distances or treatment density within fixed areas; point estimates are likely to be very sensitive to the definition of distance or area.

6.3.2 Publication bias

Opportunistic measurement of spillovers (as opposed to pre-specified measurement) may increase the chance of publication bias. In 33 out of the 54 included studies, spillover measurement did not appear to be part of the design of the original study measuring intervention impact. In these studies, it is possible that investigators discovered evidence of spillovers while measuring the direct effects of a program. Investigators may not have noticed an absence of spillovers or may have chosen not to publish findings related to spillover absence. Since efforts to estimate health spillovers outside of the vaccine literature have only recently become more common, it is not surprising to find that studies that found statistically significant spillovers were more likely to be mentioned in publications. Because many of the

included studies appeared to measure spillovers opportunistically, the designs used were in some cases suboptimal for spillover measurement. Such studies may be more prone to bias than studies with pre-specified spillover measurement.

6.4. Limitations and Potential Biases in the Review Process

We excluded studies from high-income countries from this review since our focus was on interventions relevant to populations in low- and middle-income countries. This focus was a requirement of our funders. However, there are relevant papers measuring health spillovers from high income countries, many of which evaluate vaccines (Piedra et al. 2005; Metlay et al. 2006; de Heer et al. 2011).

We did not include qualitative studies in this review because our review defined spillovers within the potential outcomes model, which is relevant to quantitative analyses. This restriction may have biased us towards inclusion of studies evaluating certain types of interventions, outcomes, and spillovers. It is possible that studies focusing on spillovers of certain types of interventions, particularly those focusing social factors affecting health, are more likely to be evaluated qualitatively.

While we made every effort to conduct a comprehensive search for spillover articles, since the concept of spillovers is poorly indexed, it is possible that we missed some relevant articles. Greater consistency in the use of terms that describe spillovers would improve future efforts to identify relevant papers by searching electronic databases. For example, some relevant papers which may have been eligible came to our attention during this review but were not identified through our search process, so we did not include them in the review (Duflo 2000; Barham 2012; Spears 2013; Baird et al. 2014).

Some of the databases we searched (e.g. Google Scholar) do not allow for repeatable searches. Thus, our complete search set cannot be fully replicated. Some titles and abstracts could only be reviewed by one team member. It is possible that there was misclassification that would have been prevented by duplicate review. Multiple team members assessed the risk of bias in a subset of studies, thus some misclassification may also have occurred in this assessment.

For the parameters for which we standardized point estimates, the information needed to convert standard errors from the additive to the relative scale was not available in the included studies, so our comparison of estimates across studies did not take precision into account. More standardized, systematic reporting across disciplines, particularly in the social sciences, would increase comparability across studies and allow for more careful assessment of risk of bias in studies (Miguel et al. 2014).

7. Conclusions

A wide range of methods with varying levels of quality has been used across academic disciplines to estimate a variety of health spillover parameters in low- and middle-income countries. The strongest evidence for spillovers was present in studies evaluating vaccines and mass drug administration for control of parasites. Future studies would benefit from incorporation of spillover measurement in the design phase and clear descriptions of methods used to estimate spillovers, ideally in a registered protocol. Rigorous estimates of spillover effects improve our understanding of the population-level impact and cost-effectiveness of interventions.

7.1. Implications for Practice and Policy

If positive spillovers are present, studies that only estimate effects on intervention recipients will underestimate the effectiveness of the intervention. In addition, cost-effectiveness calculations that exclude positive spillovers will underestimate intervention benefits. Thus, from a policy perspective, careful assessment of spillovers in future impact evaluations will allow for more comprehensive and accurate assessments of which programs yield the greatest health impacts and have the best cost-effectiveness.

Because of the low number of studies producing high quality evidence in this systematic review, we do not recommend any policies specific to particular interventions based on the results of our review. Vaccines were the only intervention assessed in included studies for which there were sufficient high quality studies to recommend further implementation because of strong evidence of spillovers. However, the strength of spillovers for many vaccines is well-documented and is already a major motivator for many governments and other organizations to implement vaccine programs. For other interventions included in this review, evidence from more rigorous evaluations in other study populations and settings would create an evidence base for decisions about future programming.

We recommend continued measurement of spillovers in future impact evaluations with improved design, and analysis and reporting (see Section 7.3). In some cases, spillover measurement can easily be incorporated into existing study designs at minimal cost. However, it is often the case that measuring spillovers rigorously requires measuring outcomes among individuals who were not targeted by a program, which may increase the cost of an impact evaluation since such individuals are not typically enrolled. Thus, funders of impact evaluations may consider whether it is appropriate to provide additional funding for assessment of spillovers for interventions when there is strong theory to suggest that they might be present.

7.2. Implications for Research

7.2.1 Pre-specification of spillover parameters, design, and analysis methods

The majority of included studies did not clearly define the spillover parameter that they estimated. We found that studies that did not pre-specify spillover estimation were more likely to have a moderate or high risk of bias; this finding underscores the importance of pre-specification. We recommend that in future studies measuring spillovers, investigators provide a

clear definition of the spillover parameter, ideally in a publicly available protocol registered prior to data collection or analysis (Miguel et al. 2014). We also recommend pre-specification of a number of other elements of study design, such as the scale of spillovers expected and the hypothesized mechanism of spillovers. Another important element that applies to some spillover studies is the definition of the area within which spillovers are to be estimated. Several of the commonly estimated spillover parameters (e.g. treatment coverage mean, spillovers conditional on treatment density) require definition of the area in which treatment density is estimated. We found that the definition of area varied widely from study to study, and frequently an explicit rationale was not given for area definitions. Since spillover magnitude is likely to be very sensitive to the area definition (Openshaw 1984), we recommend that future studies pre-define areas within which spillovers are estimated and provide a rationale for their choice of area definition.

At the reporting stage, we recommend that investigators that measure spillovers in future impact evaluations use the checklist we developed building on the CONSORT framework to ensure thorough reporting of spillovers, increase standardization across disciplines, and allow for greater comparability of spillover findings.

7.2.2 Choice of spillover parameter and design

The choice of which spillover parameter to estimate depends on a variety of factors including the level of intervention delivery (individual vs. group), the hypothesized mechanism of spillover, and the information available to investigators (e.g., certain social network parameters may require a near full census of the study population). As for all studies, regardless of the parameter chosen, studies that estimate spillovers and wish to make causal inferences must employ rigorous designs that randomize or approximate randomization (e.g., through multivariate matching) and minimize contamination of the control group. In the cases in which contamination of the control group occurs, in general, spillover estimates can be considered lower bounds of the true spillover effect under certain assumptions (discussed in Section 3.2.1). There is, however, one parameter that we do not recommend estimating in future studies: the treatment coverage mean/effect. This parameter is limited to ecologic inferences and may be subject to the ecologic fallacy (Morgenstern 1982). We recommend that studies that wish to rigorously estimate spillovers utilize the other parameters discussed in this review, which require individual level measurements but provide stronger evidence of spillovers.

7.2.3 Statistical power and spillovers

We recommend that future studies conduct sample size calculations to ensure that statistical power is sufficient to detect spillover effects. Certain studies that leverage cluster-randomized or individually-randomized designs with minor design modifications to measure spillovers do not require specialized sample size formulas. However, since it is typically the case that investigators assume spillover effects will be smaller than direct or total effects, it is prudent to include conservative estimates of spillover effects in these sample size calculations. Furthermore, if spillover effects are expected to be negative, even larger sample may be needed to have sufficient statistical power to detect them. For non-standard designs, such as

the double-randomized design, simulations can be used to estimate the required sample size (Arnold et al. 2011).

7.2.4 Unanticipated spillover findings

While we recommend pre-specification of spillover measurement in protocol registries, we recognize that in some instances, study investigators may not anticipate spillovers, particularly for interventions for which spillovers are unexpected or have never previously been documented. Indeed, many of the studies included in this review that are the first to document spillovers of an intervention do so without prior specification (e.g. spillovers of sari cloth filtration (Huq et al. 2010)). These studies serve as an important foundation for future studies to assess whether spillovers are present using more rigorous methods and in other study populations. In such cases, reporting the spillover effects is still valuable with the appropriate caveat that they were not pre-specified because those reports will help generate new hypotheses and/or motivate additional, confirmatory studies.

7.3. Recommendations

In this section, we make recommendations for the design and analysis of studies that wish to measure spillovers as well as recommendations for reporting spillovers.

7.3.1 Design and analysis recommendations

- Pre-specify spillover estimation when developing study designs or analysis plans. This includes mentioning spillover estimation methods during protocol registration.
- Pre-specify the specific spillover parameter(s) to be estimated.
- If the spillover parameter incorporates measurement within specific distances or areas, pre-specify these distances or areas and provide a rationale for them. If measures of treatment or outcome density are to be used, pre-specify the definition of areas or the algorithm used to define areas within which density is measured with as much detail as possible. For example, describe the specific distances in which measurement will take place or describe whether measurement will occur within quantiles of the observed distance distribution.
- Define spillover parameters that include contrasts that are likely to be policy relevant. For example, in estimating social network spillovers, estimate the difference in outcomes for a realistic number of social network links compared to no social network links.
- Pre-specify the scale at which spillovers are expected and the hypothesized mechanism(s) of spillover.
- If the study protocol is registered, use the term “spillovers” or “indirect effects” to refer to spillovers in the protocol because these are the most commonly used terms in the literature (Table 10), and they provide a direct link to the theoretical literature on this topic.

- To estimate within-cluster spillovers, utilize a double-randomized design. If it is only possible to utilize a cluster-randomized design, consider using multivariate matching techniques to match untreated individuals in the control clusters to untreated individuals in the treatment clusters. This will ensure internal validity but may decrease external validity in some cases.
- If a clustered study design is used, build in buffer zones between treated and control units in order to prevent contamination and ensure that there is a valid control group to serve as a counterfactual.

7.3.2 Reporting recommendations for studies estimating spillovers

We have organized the following recommendations within the headings of the CONSORT checklist for reporting of randomized trials (Schulz et al. 2010).

Title and abstract

- If spillovers were measured as a primary outcome of a study, mention them in the title and/or abstract. Use the term “spillovers” or “indirect effects” to refer to spillovers.

Introduction

<u>Background and objectives</u>

- Use the term “spillovers” or “indirect effects” to refer to spillovers.

Methods

<u>Design</u>

- Indicate whether spillover estimation was pre-specified.
- Describe whether buffers existed between treatment and control units, whether in physical or social distance.
- If treatment or outcome density was measured within areas, describe the rationale for and method of defining these areas.
- Describe the scale on which spillovers are expected (e.g. household, village, etc.).
- For study designs used to estimate spillovers other than the double-randomized or the cluster-randomized design, provide a clear description of the assumptions required to estimate valid statistical parameters if SUTVA is violated.

<u>Participants</u>

- Provide a clear description of the rationale for treatment eligibility criteria.
- State whether individuals enrolled to measure spillovers were eligible for the treatment or not.

<u>Interventions</u>

- Provide a clear description of how treatment was allocated to groups and individuals.
- State whether the level of treatment allocation was chosen in order to measure spillovers.
- Describe whether untreated individuals in treated areas were randomly assigned to not receive treatment, if they opted out of treatment, if they were ineligible for treatment, or if there were other reasons they were not treated.
- Describe the mechanism of spillovers hypothesized and assessed for each treatment.
- Describe whether a buffer zone was created between treatment and control units.

Outcomes

- If outcomes measured to estimate direct or total effects differed from outcomes measured to estimate spillover effects, provide a rationale for the difference.

Sample size

- Describe any calculations conducted to determine the sample size needed to estimate spillover parameters. If none, state that none were conducted.

Statistical methods

- Define the specific spillover parameter(s) estimated for each intervention.
- Describe the design and statistical analysis methods used to identify spillover effects.
- Describe any assumptions underlying statistical methods used to estimate spillovers.
- Describe whether any unplanned analyses were conducted to estimate spillovers. These may include subgroup analyses or analyses with alternative definitions of areas in which treatment coverage or density was estimated.

Results

Participant flow

- Provide the number of clusters allocated to treatment and control.
- Provide the number of individuals eligible to receive treatment in treated clusters.
- Provide the number of individuals allocated to treatment within treatment clusters, allocated to not receive treatment within treated clusters, and allocated to control clusters.
- Provide the number of individuals that received and did not treatment within treatment and control clusters.
- For spillovers measured within clusters, provide information about the proportion of individuals receiving treatment within each cluster.

- If measurement occurred in buffer zones between treatment and control clusters, provide the number of individuals who did and did not receive treatment in buffer zones.

Recruitment

- If dates of data collection for spillover measures differed from dates for total or direct effect measures, explain the discrepancy.

Outcomes and estimation

- Clearly label which results estimate each spillover parameter.
- In tables and figures, clearly indicate whether parameters were estimated among individuals allocated to treatment vs. those that received treatment (i.e., indicate whether an intention-to-treat vs. treatment-on-treated analysis was conducted).
- If multiple spillover mechanisms were hypothesized, label results according to the hypothesized spillover mechanism.
- Present total effects for comparable population subgroups to allow for assessment of the proportion of the total effect attributable to spillovers.
- If direct or total effects are estimated in subgroups, present spillover estimates in these same subgroups to allow for direct comparison.
- Report whether there was any evidence that untreated individuals in the treatment or control group were exposed to treatment (e.g., if untreated individuals had heard of the intervention or knew individuals who received it).
- Describe any evidence of contamination of the control group.

Discussion

- Present evidence supporting the proposed mechanism of spillover. Such evidence may or may not have been collected in the study but contributes substantially to the overall understanding and credibility of the results.

Limitations

- Discuss any potential biases that may be present for spillover parameters. Discuss whether these biases may also be present for direct or total effect parameters. This includes contamination of the control group.
- Articulate whether any analyses conducted to estimate spillovers were not pre-specified.

References

References to Included Studies

Ali M, Emch M, von Seidlein L, et al (2005) Herd immunity conferred by killed oral cholera vaccines in Bangladesh: a reanalysis. Lancet 366:44–49. doi: 10.1016/S0140-6736(05)66550-6

Ali M, Emch M, Yunus M, et al (2008) Vaccine Protection of Bangladeshi Infants and Young Children Against Cholera: Implications for Vaccine Deployment and Person-to-Person Transmission. Pediatr Infect Dis J 27:33–37. doi: 10.1097/INF.0b013e318149dffd

Ali M, Sur D, You YA, et al (2013) Herd protection by a bivalent-killed-whole-cell oral cholera vaccine in the slums of Kolkata, India. Clin Infect Dis cit009. doi: 10.1093/cid/cit009

Avitabile C (2012) Spillover Effects in Healthcare Programs: Evidence on Social Norms and Information Sharing. Inter-American Development Bank

Azad K, Barnett S, Banerjee B, et al (2010) Effect of scaling up women's groups on birth outcomes in three rural districts in Bangladesh: a cluster-randomised controlled trial. The Lancet 375:1193–1202. doi: 10.1016/S0140-6736(10)60142-0

Baird S, De Hoop J, Özler B (2013a) Income shocks and adolescent mental health. J Hum Resour 48:370–403.

Baird S, Hicks JH, Kremer M, Miguel E (2013b) Worms at Work: Public Finance Implications of a Child Health Investment.

Banerjee AV, Duflo E, Glennerster R, Kothari D (2010) Improving immunisation coverage in rural India: clustered randomised controlled evaluation of immunisation campaigns with and without incentives. BMJ 340:c2220–c2220. doi: 10.1136/bmj.c2220

Baptista PN, Magalhães V, Rodrigues LC, et al (2006) Pertussis vaccine effectiveness in reducing clinical disease, transmissibility and proportion of cases with a positive culture after household exposure in Brazil. Pediatr Infect Dis J 25:844–846. doi: 10.1097/01.inf.0000232642.25495.95

Bhattacharya D, Dupas P, Kanaya S (2013) Estimating the Impact of Means-tested Subsidies under Treatment Externalities with Application to Anti-Malarial Bednets. National Bureau of Economic Research

Björkman M, Svensson J (2009) Power to the people: evidence from a randomized field experiment on community-based monitoring in Uganda. Q J Econ 124:735–769.

Buttenheim A, Alderman H, Friedman J (2011) Impact evaluation of school feeding programmes in Lao People's Democratic Republic. J Dev Eff 3:520–542. doi: 10.1080/19439342.2011.634511

Chaudhuri A (2005) Intra-Household Spillover Effects of a Maternal and Child Health Program: Evidence from Rural Bangladesh.

Chen W-J, Moulton LH, Saha SK, et al (2014) Estimation of the herd protection of Haemophilus influenzae type b conjugate vaccine against radiologically confirmed pneumonia in children under 2 years old in Dhaka, Bangladesh. Vaccine 32:944–948. doi: 10.1016/j.vaccine.2013.12.052

Chidambaram JD, Melese M, Alemayehu W, et al (2004) Mass Antibiotic Treatment and Community Protection in Trachoma Control Programs. Clin Infect Dis 39:e95–e97. doi: 10.1086/424747

Chong A, Gonzalez-Navarro M, Karlan D, Valdivia M (2013) Effectiveness and Spillovers of Online Sex Education: Evidence from a Randomized Evaluation in Colombian Public Schools. National Bureau of Economic Research

Contreras D, Maitra P (2013) Health Spillover Effects of a Conditional Cash Transfer Program. Monash University, Department of Economics

Cooper E, Fitch L (1983) Pertussis: Herd Immunity and Vaccination Coverage in St. Lucia. The Lancet 322:1129–1132. doi: 10.1016/S0140-6736(83)90637-2

Dupas P (2006) Relative Risks and the Market for Sex: Teenagers, Sugar Daddies and HIV in Kenya. http://mpra.ub.uni-muenchen.de/248/. Accessed 13 Jul 2014

Egere U, Townend J, Roca A, et al (2012) Indirect effect of 7-valent pneumococcal conjugate vaccine on pneumococcal carriage in newborns in rural Gambia: a randomised controlled trial. PloS One 7:e49143. doi: 10.1371/journal.pone.0049143

Emch M, Ali M, Park J-K, et al (2006) Relationship between neighbourhood-level killed oral cholera vaccine coverage and protective efficacy: evidence for herd immunity. Int J Epidemiol 35:1044–1050. doi: 10.1093/ije/dyl100

Emch M, Ali M, Root ED, Yunus M (2009) Spatial and environmental connectivity analysis in a cholera vaccine trial. Soc Sci Med 68:631–637. doi: 10.1016/j.socscimed.2008.11.025

Fitzsimons E, Malde B, Mesnard A, Vera-Hernández M (2012) Household responses to information on child nutrition: experimental evidence from Malawi.

Forleo-Neto E, Oliveira CF de, Maluf EMCP, et al (1999) Decreased Point Prevalence of Haemophilus influenzae Type b (Hib) Oropharyngeal Colonization by Mass Immunization of Brazilian Children Less Than 5 Years Old with Hib Polyribosylribitol Phosphate Polysaccharide—Tetanus Toxoid Conjugate Vaccine in Combination with Diphtheria-Tetanus Toxoids—Pertussis Vaccine. J Infect Dis 180:1153–1158. doi: 10.1086/315018

German D, Sutcliffe CG, Sirirojn B, et al (2012) Unanticipated Effect of a Randomized Peer Network Intervention on Depressive Symptoms Among Young Methamphetamine Users in Thailand. J Community Psychol 40:799–813. doi: 10.1002/jcop.21488

Godlonton S, Thornton R (2012) Peer effects in learning HIV results. J Dev Econ 97:118–129.

Haile M, Tadesse Z, Gebreselassie S, et al (2013) The Association between Latrine Use and Trachoma: A Secondary Cohort Analysis from a Randomized Clinical Trial. Am J Trop Med Hyg 89:717–720. doi: 10.4269/ajtmh.13-0299

Hammitt LL, Akech DO, Morpeth SC, et al (2014) Population effect of 10-valent pneumococcal conjugate vaccine on nasopharyngeal carriage of Streptococcus pneumoniae and non-typeable Haemophilus influenzae in Kilifi, Kenya: findings from cross-sectional carriage studies. Lancet Glob Health 2:e397–e405.

Handa S, Huerta M-C, Perez R, Straffon B (2001) Poverty, inequality, and spillover in Mexico's education, health, and nutrition program.

Hawley WA, Phillips-Howard PA, Kuile FOT, et al (2003) Community-Wide Effects of Permethrin-Treated Bed Nets on Child Mortality and Malaria Morbidity in Western Kenya. Am J Trop Med Hyg 68:121–127.

House JI, Ayele B, Porco TC, et al (2009) Assessment of herd protection against trachoma due to repeated mass antibiotic distributions: a cluster-randomised trial. The Lancet 373:1111–1118. doi: 10.1016/S0140-6736(09)60323-8

Huq A, Yunus M, Sohel SS, et al (2010) Simple Sari Cloth Filtration of Water Is Sustainable and Continues To Protect Villagers from Cholera in Matlab, Bangladesh. mBio 1:e00034–10. doi: 10.1128/mBio.00034-10

Janssens W (2005) Measuring externalities in program evaluation. Tinbergen Institute Discussion Paper

Joshi S, Schultz TP (2013) Family Planning and Women's and Children's Health: Long-Term Consequences of an Outreach Program in Matlab, Bangladesh. Demography 50:149–180.

Kazianga H, de Walque D, Alderman H (2009) School Feeding Programs and th e Nutrition of Siblings: Evidence from a Randomized Trial in Rural Burkina Faso. Oklahoma State University, Department of Economics and Legal Studies in Business

Khan MI, Soofi SB, Ochiai RL, et al (2012) Effectiveness of Vi capsular polysaccharide typhoid vaccine among children: A cluster randomized trial in Karachi, Pakistan. Vaccine 30:5389–5395. doi: 10.1016/j.vaccine.2012.06.015

Khatib AM, Ali M, von Seidlein L, et al (2012) Effectiveness of an oral cholera vaccine in Zanzibar: findings from a mass vaccination campaign and observational cohort study. Lancet Infect Dis 12:837–844. doi: 10.1016/S1473-3099(12)70196-2

Kremer M, Miguel E (2007) The illusion of sustainability. Q J Econ 122:1007–1065.

Miguel E, Kremer M (2004) Worms: Identifying Impacts on Education and Health in the Presence of Treatment Externalities. Econometrica 72:159–217. doi: 10.1111/j.1468-0262.2004.00481.x

Ozier O (2014) Exploiting Externalities to Estimate the Long-term Benefits of Early Childhood Deworming.

Paul JR, Horstmann DM, Riordan JT, et al (1962) An oral poliovirus vaccine trial in Costa Rica. Bull World Health Organ 26:311–329.

Perez-Heydrich C, Hudgens MG, Halloran ME, et al (2014) Assessing effects of cholera vaccination in the presence of interference.

Préziosi M-P, Halloran ME (2003) Effects of pertussis vaccination on transmission: vaccine efficacy for infectiousness. Vaccine 21:1853–1861. doi: 10.1016/S0264-410X(03)00007-0

Ribas RP, Soares FV, Teixeira CG, et al (2011) Externality and behavioural change effects of a non-randomised CCT programme: Heterogeneous impact on the demand for health and education. Working Paper, International Policy Centre for Inclusive Growth

Roca A, Dione MM, Bojang A, et al (2013) Nasopharyngeal carriage of pneumococci four years after community-wide vaccination with PCV-7 in The Gambia: long-term evaluation of a cluster randomized trial.

Roca A, Hill PC, Townend J, et al (2011) Effects of Community-Wide Vaccination with PCV-7 on Pneumococcal Nasopharyngeal Carriage in The Gambia: A Cluster-Randomized Trial. PLoS Med 8:e1001107. doi: 10.1371/journal.pmed.1001107

Root ED, Giebultowicz S, Ali M, et al (2011) The Role of Vaccine Coverage within Social Networks in Cholera Vaccine Efficacy. PLoS ONE 6:e22971.

Root ED, Lucero M, Nohynek H, et al (2014) Distance to health services affects local-level vaccine efficacy for pneumococcal conjugate vaccine (PCV) among rural Filipino children. Proc Natl Acad Sci 111:3520–3525. doi: 10.1073/pnas.1313748111

Shekhawat N, Mkocha H, Munoz B, et al (2014) Cohort and Age Effects of Mass Drug Administration on Prevalence of Trachoma: A Longitudinal Study in Rural Tanzania. Invest Ophthalmol Vis Sci 55:2307–2314. doi: 10.1167/iovs.13-12701

Singh P. Spillovers in learning and behavior: Evidence from a nutritional information campaign in urban slums.

Sur D, Ochiai RL, Bhattacharya SK, et al (2009) A Cluster-Randomized Effectiveness Trial of Vi Typhoid Vaccine in India. N Engl J Med 361:335–344. doi: 10.1056/NEJMoa0807521

Tontarawongsa C, Mahajan A, Tarozzi A (2011) (Limited) Diffusion of Health-protecting Behaviors: Evidence from Non-beneficiaries of a Public Health Program in Orissa (India).

Ziegelhöfer Z (2012) Down with diarrhea: Using fuzzy regression discontinuity design to link communal water supply with health. Graduate Institute of International and Development Studies Working Paper

Zivin JG, Thirumurthy H, Goldstein M (2009) AIDS treatment and intrahousehold resource allocation: Children's nutrition and schooling in Kenya. J Public Econ 93:1008–1015.

References to Excluded Studies

See the supplementary spreadsheet.

References to Studies Awaiting Classification

None

References to Ongoing Studies

None

Additional References

Aiken A, Davey C, Hargreaves J, Hayes R (2015) Reanalysis of health and educational impacts of a school-based deworming program in western Kenya Part 1: pure replication. International Initiative for Impact Evaluation

Anderson RM, May RM (1979) Population biology of infectious diseases: Part I. Nature 280:361–367. doi: 10.1038/280361a0

Angelucci M, Maro VD (2010) Program Evaluation and Spillover Effects. Inter-American Development Bank

Angrist JD, Imbens GW, Rubin DB (1996) Identification of Causal Effects Using Instrumental Variables. J Am Stat Assoc 91:444–455. doi: 10.1080/01621459.1996.10476902

Arnold BF, Hogan DR, Colford JM, Hubbard AE (2011) Simulation methods to estimate design power: an overview for applied research. BMC Med Res Methodol 11:94. doi: 10.1186/1471-2288-11-94

Auchincloss AH, Diez Roux AV (2008) A New Tool for Epidemiology: The Usefulness of Dynamic-Agent Models in Understanding Place Effects on Health. Am J Epidemiol 168:1–8. doi: 10.1093/aje/kwn118

Baird S, Bohren A, McIntosh C, Ozler B (2014) Designing experiments to measure spillover and threshold effects. PIER Working Paper No. 14-006

Balshem H, Helfand M, Schünemann HJ, et al (2011) GRADE guidelines: 3. Rating the quality of evidence. J Clin Epidemiol 64:401–406. doi: 10.1016/j.jclinepi.2010.07.015

Bandura A (1986) Social foundations of thought and action. Prentice Hall, Englewood Cliffs, NJ

Banerjee A, Chandrasekhar AG, Duflo E, Jackson MO (2013) The Diffusion of Microfinance. Science 341:1236498. doi: 10.1126/science.1236498

Banerjee AV (1992) A Simple Model of Herd Behavior. Q J Econ 107:797–817. doi: 10.2307/2118364

Barham T (2012) Enhancing Cognitive Functioning: Medium-Term Effects of a Health and Family Planning Program in Matlab. Am Econ J Appl Econ 4:245–273. doi: 10.1257/app.4.1.245

Barnes JA (1954) Class and Committees in a Norwegian Island Parish. Plenum

Berkman LF, Syme SL (1979) Social networks, host resistance, and mortality: a nine-year follow-up study of Alameda County residents. Am J Epidemiol 109:186–204.

Berk R, Freedman D (2010) Statistical assumptions as empirical commitments. In: Collier D, Sekhon J, Stark P (eds) Statistical Models and Causal Inference. pp 23–43

Bikhchandani S, Hirshleifer D, Welch I (1992) A Theory of Fads, Fashion, Custom, and Cultural Change as Informational Cascades. J Polit Econ 100:992–1026.

Bor J, Moscoe E, Mutevedzi P, et al (2014) Regression Discontinuity Designs in Epidemiology. Epidemiol Camb Mass 25:729–737. doi: 10.1097/EDE.0000000000000138

Bott E (1957) Family and Social Network: Roles, Norms, and External Relationships in Ordinary Families. Tavistock

Chabot I, Goetghebeur MM, Grégoire J-P (2004) The societal value of universal childhood vaccination. Vaccine 22:1992–2005. doi: 10.1016/j.vaccine.2003.10.027

Clemens J, Shin S, Ali M (2011) New approaches to the assessment of vaccine herd protection in clinical trials. Lancet Infect Dis 11:482–487. doi: 10.1016/S1473-3099(10)70318-2

Coalition for Evidence-Based Policy (2010) Checklist For Reviewing a Randomized Controlled Trial of a Social Program or Project, To Assess Whether It Produced Valid Evidence. http://coalition4evidence.org/wp-content/uploads/Checklist-For-Reviewing-a-RCT-Jan10.pdf.

Colford J (2015) Spillover Effects of Water, Sanitation, and Hygiene Interventions on Child Health. https://clinicaltrials.gov/ct2/show/NCT02396407.

Cox DR (1958) Planning of experiments. Wiley, Oxford, England

Curns AT, Steiner CA, Barrett M, et al (2010) Reduction in Acute Gastroenteritis Hospitalizations among US Children After Introduction of Rotavirus Vaccine: Analysis of Hospital Discharge Data from 18 US States. J Infect Dis 201:1617–1624. doi: 10.1086/652403

Davey C, Aiken A, Hayes R, Hargreaves J (2014) Reanalysis of health and educational impacts of a school-based deworming program in western Kenya Part 2: alternative analyses. International Initiative for Impact Evaluation

de Heer HD, Koehly L, Pederson R, Morera O (2011) Effectiveness and Spillover of an After-School Health Promotion Program for Hispanic Elementary School Children. Am J Public Health 101:1907–1913. doi: 10.2105/AJPH.2011.300177

Diamond A, Sekhon JS (2013) Genetic Matching for Estimating Causal Effects: A General Multivariate Matching Method for Achieving Balance in Observational Studies. Rev Econ Stat 95:932–945. doi: 10.1162/REST_a_00318

do Carmo GMI, Yen C, Cortes J, et al (2011) Decline in Diarrhea Mortality and Admissions after Routine Childhood Rotavirus Immunization in Brazil: A Time-Series Analysis. PLoS Med 8:e1001024. doi: 10.1371/journal.pmed.1001024

Duflo E (2000) Child Health and Household Resources in South Africa: Evidence from the Old Age Pension Program. Am Econ Rev 90:393–398.

Dupas P (2014) Short-Run Subsidies and Long-Run Adoption of New Health Products: Evidence From a Field Experiment. Econometrica 82:197–228. doi: 10.3982/ECTA9508

Dybvig PH, Spatt CS (1983) Adoption externalities as public goods. J Public Econ 20:231–247.

Effective Practice and Organisation of Care (EPOC) Group (2009) Risk of bias for studies with a separate control group. http://epoc.cochrane.org/sites/epoc.cochrane.org/files/uploads/Risk%20of%20Bias%2005-01-2009.doc.

Ellison G, Fudenberg D (1995) Word-of-Mouth Communication and Social Learning. Q J Econ 110:93–125. doi: 10.2307/2118512

Everitt B, Landau S, Leese M, Stahl D (2011) Cluster Analysis, 5th edition. Wiley, Chichester

Fine PE (1993) Herd immunity: history, theory, practice. Epidemiol Rev 15:265–302.

Fitzsimons E, Malde B, Mesnard A, Vera-Hernández M (2012) Household responses to information on child nutrition: experimental evidence from Malawi.

Fox JP, Elveback L, Scott W, et al (1995) Herd immunity: basic concept and relevance to public health immunization practices. 1971. Am J Epidemiol 141:187–197; discussion 185–186.

Galea S, Riddle M, Kaplan GA (2010) Causal thinking and complex system approaches in epidemiology. Int J Epidemiol 39:97–106. doi: 10.1093/ije/dyp296

Gertler PJ (2010) Impact Evaluation in Practice. http://publications.worldbank.org/index.php?main_page=product_info&products_id=23915. Accessed 6 Jan 2012

Giebultowicz S, Ali M, Yunus M, Emch M (2011) The simultaneous effects of spatial and social networks on cholera transmission. Interdiscip Perspect Infect Dis 2011:1–6. doi: 10.1155/2011/604372

Greenland S (2000) An introduction to instrumental variables for epidemiologists. Int J Epidemiol 29:722–729. doi: 10.1093/ije/29.4.722

Grijalva CG, Nuorti JP, Arbogast PG, et al (2007) Decline in pneumonia admissions after routine childhood immunisation with pneumococcal conjugate vaccine in the USA: a time-series analysis. The Lancet 369:1179–1186. doi: 10.1016/S0140-6736(07)60564-9

Guyatt GH, Oxman AD, Vist GE, et al (2008) GRADE: an emerging consensus on rating quality of evidence and strength of recommendations. BMJ 336:924–926. doi: 10.1136/bmj.39489.470347.AD

Haider M, Kreps G (2004) Forty Years of Diffusion of Innovations: Utility and Value in Public Health. J Health Commun 9:3–11. doi: 10.1080/10810730490271430

Halloran ME (2012) The Minicommunity Design to Assess Indirect Effects of Vaccination. Epidemiol Methods 1:83–105.

Halloran ME, Haber M, Longini IM Jr, Struchiner CJ (1991) Direct and indirect effects in vaccine efficacy and effectiveness. Am J Epidemiol 133:323–331.

Halloran ME, Longini IM, Struchiner CJ (2010) Design and analysis of vaccine studies. Springer

Halloran ME, Struchiner CJ (1995) Causal Inference in Infectious Diseases. Epidemiology 6:142–151.

Halloran ME, Struchiner CJ (1991) Study Designs for Dependent Happenings. Epidemiology 2:331–338. doi: 10.2307/20065696

Hayes R, Alexander ND, Bennett S, Cousens SN (2000) Design and analysis issues in cluster-randomized trials of interventions against infectious diseases. Stat Methods Med Res 9:95–116.

Hicks J, Kremer M, Miguel E (2014) Estimating deworming school participation impacts and externalities in Kenya: A Comment on Aiken et al. (2014). International Initiative for Impact Evaluation

Higgins J, Greene S (2011) Cochrane Handbook for Systematic Reviews of Interventions, 5.1.0 edn. Cochrane Collaboration

Holland P (1986) Statistics and Causal Inference. J Am Stat Assoc 81:945–960.

House JI, Ayele B, Porco TC, et al (2009) Assessment of herd protection against trachoma due to repeated mass antibiotic distributions: a cluster-randomised trial. The Lancet 373:1111–1118. doi: 10.1016/S0140-6736(09)60323-8

Hudgens MG, Halloran ME (2008) Toward Causal Inference With Interference. J Am Stat Assoc 103:832–842. doi: 10.1198/016214508000000292

Imbens G, Lemieux T (2007) Regression Discontinuity Designs: A Guide to Practice. Natl Bur Econ Res Tech Work Pap Ser 337.

John TJ, Samuel R (2000) Herd immunity and herd effect: new insights and definitions. Eur J Epidemiol 16:601–606. doi: 10.1023/A:1007626510002

Kermack WO, McKendrick AG (1991) Contributions to the mathematical theory of epidemics--II. The problem of endemicity.1932. Bull Math Biol 53:57–87.

Lee DS, Lemieux T (2010) Regression Discontinuity Designs in Economics. J Econ Lit 48:281–355. doi: 10.1257/jel.48.2.281

Little RJ, Rubin DB (2000) Causal Effects in Clinical and Epidemiological Studies Via Potential Outcomes: Concepts and Analytical Approaches. Annu Rev Public Health 21:121–145. doi: 10.1146/annurev.publhealth.21.1.121

Longini IM, Koopman JS, Haber M, Cotsonis GA (1988) Statistical Inference for Infectious Diseases Risk-Specific Household and Community Transmission Parameters. Am J Epidemiol 128:845–859.

Longini IM, Sagatelian K, Rida WN, Halloran ME (1998) Optimal vaccine trial design when estimating vaccine efficacy for susceptibility and infectiousness from multiple populations. Stat Med 17:1121–1136. doi: 10.1002/(SICI)1097-0258(19980530)17:10<1121::AID-SIM824>3.0.CO;2-E

Manski CF (1993) Identification of Endogenous Social Effects: The Reflection Problem. Rev Econ Stud 60:531–542. doi: 10.2307/2298123

Marsden PV (2006) Network methods in social epidemiology. In: Oakes JM, Kaufman JS (eds) Methods in social epidemiology. Jossey-Bass,

Metlay JP, Fishman NO, Joffe M, Edelstein PH (2006) Impact of pediatric vaccination with pneumococcal conjugate vaccine on the risk of bacteremic pneumococcal pneumonia in adults. Vaccine 24:468–475. doi: 10.1016/j.vaccine.2005.07.095

Miguel E, Camerer C, Casey K, et al (2014) Promoting Transparency in Social Science Research. Science 343:30–31. doi: 10.1126/science.1245317

Morgenstern H (1982) Uses of ecologic analysis in epidemiologic research. Am J Public Health 72:1336–1344. doi: 10.2105/AJPH.72.12.1336

Newman MEJ (2002) Spread of epidemic disease on networks. Phys Rev E 66:016128. doi: 10.1103/PhysRevE.66.016128

Neyman J, Dabrowska DM, Speed TP (1990) On the Application of Probability Theory to Agricultural Experiments. Essay on Principles. Section 9. Stat Sci 5:465–472. doi: 10.1214/ss/1177012031

Oakes JM, Kaufman JS (2006) Methods in social epidemiology. John Wiley & Sons

O'Malley AJ, Marsden PV (2008) The analysis of social networks. Health Serv Outcomes Res Methodol 8:222–269. doi: 10.1007/s10742-008-0041-z

Openshaw S (1984) Ecological fallacies and the analysis of areal census data. Environ Plan A 16:17–31.

Perkins RB, Langrish S, Stern LJ, Simon CJ (2007) A community-based education program about cervical cancer improves knowledge and screening behavior in Honduran women. Rev Panam Salud Pública 22:187–193. doi: 10.1590/S1020-49892007000800005

Piedra PA, Gaglani MJ, Kozinetz CA, et al (2005) Herd immunity in adults against influenza-related illnesses with use of the trivalent-live attenuated influenza vaccine (CAIV-T) in children. Vaccine 23:1540–1548. doi: 10.1016/j.vaccine.2004.09.025

Quian J, Rüttimann R, Romero C, et al (2008) Impact of universal varicella vaccination on 1-year-olds in Uruguay: 1997–2005. Arch Dis Child 93:845–850. doi: 10.1136/adc.2007.126243

Rogers EM (2010) Diffusion of Innovations, 4th Edition. Simon and Schuster

Rosenbaum PR (2007) Interference Between Units in Randomized Experiments. J Am Stat Assoc 102:191–200. doi: 10.1198/016214506000001112

Rosenbaum PR, Rubin DB (1983) The central role of the propensity score in observational studies for causal effects. Biometrika 70:41 –55. doi: 10.1093/biomet/70.1.41

Rosenbaum PR, Rubin DB (1985) The bias due to incomplete matching. Biometrics 41:103–116.

Ross R (1915) Some a priori pathometric equations. Br Med J 1:546–547.

Rubin D (1974) Estimating Causal Effects of Treatments In Randomized and Nonrandomized Studies. J Educ Psychol 66:688–701.

Rubin DB (1990) Formal mode of statistical inference for causal effects. J Stat Plan Inference 25:279–292. doi: 10.1016/0378-3758(90)90077-8

Rubin DB (1976) Inference and missing data. Biometrika 63:581–592. doi: 10.1093/biomet/63.3.581

Rubin DB (1978) Bayesian Inference for Causal Effects: The Role of Randomization. Ann Stat 6:34–58.

Schulz K, Altman D, Moher D, the CONSORT Group (2010) CONSORT 2010 Statement: updated guidelines for reporting parallel group randomised trials. BMC Med 8:18. doi: 10.1186/1741-7015-8-18

Shakya HB, Christakis NA, Fowler JH (2014) Social network predictors of latrine ownership.

Sinclair B, McConnell M, Green DP (2012) Detecting spillover effects: Design and analysis of multilevel experiments. Am J Polit Sci 56:1055–1069.

Smith KP, Christakis NA (2008) Social Networks and Health. Annu Rev Sociol 34:405–429. doi: 10.1146/annurev.soc.34.040507.134601

Spears D (2013) Essays in the Economics of Sanitation and Human Capital in Developing Countries.

Tchetgen EJT, VanderWeele TJ (2012) On causal inference in the presence of interference. Stat Methods Med Res 21:55–75. doi: 10.1177/0962280210386779

van der Laan M (2012) Causal Inference for Networks.

VanderWeele TJ, Tchetgen Tchetgen EJ (2011) Effect partitioning under interference in two-stage randomized vaccine trials. Stat Probab Lett 81:861–869. doi: 10.1016/j.spl.2011.02.019

VanderWeele T, Tchetgen Tchetgen E, Halloran M (2012) Components of the Indirect Effect in Vaccine Trials: Identification of Contagion and Infectiousness Effects. Epidemiology 23:751–761.

Vermeersch C, Kremer M (2004) School Meals, Educational Achievement, and School Competition: Evidence from a Randomized Evaluation.

Wamai RG, Ayissi CA, Oduwo GO, et al (2012) Assessing the Effectiveness of a Community-Based Sensitization Strategy in Creating Awareness About HPV, Cervical Cancer and HPV

Vaccine Among Parents in North West Cameroon. J Community Health 37:917–926. doi: 10.1007/s10900-012-9540-5

Wells G, Shea B, O'Connell D, et al (2011) The Newcastle-Ottawa Scale (NOS) for assessing the quality of nonrandomised studies in meta-analyses. http://www.ohri.ca/programs/clinical_epidemiology/oxford.asp. Accessed 3 Mar 2013

World Bank (2012) How we Classify Countries | Data. http://data.worldbank.org/about/country-classifications. Accessed 1 Mar 2012

8. Plans for Updating the Review

The co-authors of this report currently do not have plans to update this review.

9. Acknowledgments

We would like to thank John Eyers for his valuable input on the search and data collection protocol. We would also like to thank David Addiss, Isha Ray, and Art Reingold, for their helpful input on the study protocol and report.

10. The Review Team and Contributions of Authors

Jade Benjamin-Chung, Ph.D.	Led the systematic review process, including the development of the study protocol, management and participation in the search and review process, synthesis of the data, and drafting of the report
Jaynal Abedin, M.Sc.	Contributed to the development of the study protocol
David Berger	Contributed to the study protocol, reviewed of study records, appraised study quality
Ashley Clark, M.P.P M.A.-I.A.S.	Contributed to the data collection process and reviewed study records
Veronica Jimenez	Contributed to the data collection process and reviewed study records
Diana Tran	Contributed to the data collection process and reviewed study records
Eugene Konagaya	Contributed to the data collection process and reviewed study records
Lauren Falcao	Contributed to the development of the study protocol
Benjamin F. Arnold, Ph.D.	Contributed to the development of the study protocol, appraised study quality, contributed to the drafting of the report, and provided scientific guidance throughout the review process
Alan Hubbard, Ph.D.	Contributed to the development of the study protocol, contributed to the drafting of the report, and provided scientific guidance throughout the review process
Stephen P. Luby, M.D.	Contributed to the development of the study protocol, contributed to the drafting of the report, and provided scientific guidance throughout the review process
Edward Miguel, Ph.D.	Contributed to the development of the study protocol, contributed to the drafting of the report, and provided scientific guidance throughout the review process
John M. Colford, Jr., M.D. Ph.D.	Contributed to the development of the study protocol, contributed to the drafting of the report, and provided scientific guidance throughout the review process

11. Statement Concerning Conflict of Interest

We have no conflicts of interest to declare. Dr. Miguel and Dr. Luby, who co-authored studies included in this review, had no role in selecting studies for inclusion and did not participate in the classification of risk of bias criteria for any included studies. David Berger, a doctoral student under Dr. Miguel's supervision, participated in risk of bias assessment for Miguel and Kremer (2004).

12. Tables and Figures

Table 1: Primary design, scale, and mechanism for spillover parameter classes

Parameter class*	**Primary design type**	**Spillover scale†**	**Primary mechanism**
Treatment coverage mean/effect	Cluster	Varies	Geographic
Within-cluster spillover	Cluster	Small	Geographic
Distance-based spillover	Cluster	Large	Geographic
Spillovers conditional on exposure to cases	Cluster or non-cluster	Small	Geographic or social
Spillover conditional on treatment density	Cluster or non-cluster	Varies	Geographic or social
Social network spillover	Cluster or non-cluster	Varies	Social

* The parameter class is the category of spillover parameter as defined in section 2.2.2.

† The spillover scale is the approximate magnitude of distance at which spillovers are expected. For example, a spillover might be expected to occur only within households, in which case the scale would be small. If spillovers occurred throughout a city, the scale would be large.

Table 2a: Modifications to Cochrane "GRADE" tool to incorporate spillover assessment: classification of studies' underlying methodology

High quality • Double-randomized trials estimating within-cluster spillovers • Cluster-randomized trials estimating within-cluster spillovers among people who were not eligible but were highly comparable to eligible individuals • Individually randomized studies estimating spillover effects among social network members • Studies estimating spillovers conditional on treatment or outcome density in a randomized study in which treatment density is estimated over multiple distances (physical or social) and in which distance cutoffs are defined based on quantiles or other objective criteria • Household-based studies estimating vaccine efficacy parameters that match index cases with household controls
Moderate quality • Randomized studies estimating within-cluster spillovers among people who chose not to participate in the intervention (i.e., participants within clusters weren't randomized to receive treatment, so selection bias is possible in spillover effects)

• Cluster-randomized trials estimating within-cluster spillovers among people who were not eligible and were not highly comparable to eligible individuals • Observational studies estimating within-cluster spillovers • Studies estimating spillovers conditional on treatment or outcome density in a randomized study in which treatment or outcome density is estimated over only one distance level (physical or social) • Studies estimating spillovers conditional on treatment or outcome density in a randomized study in which treatment or outcome density is estimated and distance cutoffs were not based on objective criteria • Ecologic studies comparing outcomes over levels of treatment coverage in which the treatment was randomized and a possible dose-response pattern for spillovers was assessed
Low quality • Ecological studies comparing outcomes over levels of treatment coverage in which the treatment was not randomized • Ecological studies comparing outcomes over levels of treatment coverage that did not assess a possible dose-response gradient for spillover effects • Studies in which instrumental variables were the primary identification strategy but the exclusion restriction suffers from obvious violations or the instrument is not strongly associated with the treatment • Studies in which systematic differences were likely to be present between intervention and control group (e.g., a cohort study that did not use matching to make the control group comparable to the intervention group) • Studies that did not include a rigorous control group
Very low quality • Studies with any underlying methodology subject to serious additional concerns about risk of bias and the quality of evidence

Table 2b: Modifications to Cochrane "GRADE" tool to incorporate spillover assessment: factors that may increase or decrease the quality level of a body of evidence

Factors that may increase (i.e., "upgrade") the quality level of a body of evidence	**Factors that may decrease (i.e., "downgrade") the quality level of a body of evidence**
1. Large magnitude of spillover effect that is plausible relative to the size of the direct or total effect 2. All plausible confounding of the spillover effect would reduce a demonstrated effect or suggest a spurious effect when results show no effect 3. Dose-response gradient for spillover effect	1. Limitations in the design and implementation specific to spillover effects suggesting high likelihood of bias. These include the GRADE criteria as well as the following criteria specific to spillover effects: a. Contamination of the control group may have occurred or did occur b. Magnitude of spillover effect relative to direct/total effect does not seem plausible c. Spillover effects were not explicitly reported in the published manuscript 2. Indirect evidence 3. Unexplained heterogeneity or inconsistency of results 4. Imprecision of results 5. High probability of publication bias

Table 3: Characteristics of studies meeting inclusion criteria

Author & Year	Academic Discipline	Country	Intervention(s)	Primary Outcome(s)*	Primary Design	Spillover-related Parameters (Parameter # in Table 8)
Ali et al., 2005	Public Health	Bangladesh	Cholera vaccine	Cholera	Re-analysis of an individually randomized trial	Treatment coverage mean (1), Treatment coverage effect (2)
Ali et al., 2008	Public Health	Bangladesh	Cholera vaccine	Cholera	Re-analysis of an individually randomized trial	Treatment coverage effect (2)
Ali et al., 2013	Public Health	India	Cholera vaccine	Cholera	Cluster-randomized trial	Direct effect (3), Within-cluster spillover effect (4), Total effect (6), Treatment coverage mean (1), Treatment coverage effect (2)
Avitabile, 2012	Economics	Mexico	Conditional cash transfers	Screening for cervical cancer, blood sugar, and blood pressure	Re-analysis of a cluster-randomized trial	Within-cluster spillover effect among ineligibles (8), Total effect among eligibles (10)
Azad et al., 2010	Public Health	Bangladesh	Women's groups and health service strengthening	Neonatal mortality	Cluster-randomized trial	Direct effect (3), Within-cluster spillover effect conditional on exposure to treatment (12), Within-cluster spillover effect (4)
Baird et al., 2013	Economics	Kenya	Subsidized deworming	Self-reported health and BMI	Cross-sectional survey of a population that previously participated in a cluster-randomized trial	Total effect conditional on treatment density (21), Spillover effect conditional on treatment density (19)
Baird et al., 2013	Economics	Malawi	Conditional cash transfers	Psychological distress	Double-randomized trial	Within-cluster spillover effect (4), Total effect (6)
Banerjee et al., 2010	Economics	India	Vaccines (BCG, DPT, polio, measles) + incentives	Vaccine coverage	Cluster-randomized trial	Total effect (6), Spillover effect conditional on living in an untreated cluster within distance *d* to treated clusters (13)
Baptista et al., 2006	Public Health	Brazil	Pertussis vaccine	Pertussis	Case-control study	Vaccine efficacy for susceptibility (23), Vaccine efficacy for infectiousness (24)
Bhattacharya et al., 2013	Economics	Kenya	Subsidized insecticide-treated nets	Insecticide-treated net purchase	Re-analysis of an individually randomized trial	Total effect (6), Direct effect (3), Direct effect conditional on treatment density (18)

Björkman & Svensson, 2009	Economics	Uganda	Community monitoring and provision of health services	Child mortality, health service provision and utilization, child growth	Cluster-randomized trial	Total effect (6), Total effect conditional on whether treatment and control units were within distance *d* of each other (16)
Buttenheim et al., 2011	Economics	Laos	School feeding program	Child growth	Cohort study + propensity-score matching + difference-in-differences	Within-cluster spillover effect among ineligibles (8)
Chaudhuri, (year not listed)	Economics	Bangladesh	Maternal and child health program	Body mass index	Cross-sectional survey of a population previously in an unmatched cohort study	Total effect among eligibles 10), Within-cluster spillover effect among ineligibles (8)
Chen et al., 2014	Public Health	Bangladesh	*Haemophilus influenzae* type b vaccine	Pneumonia	Re-analysis of matched case-control study	Treatment coverage effect (2)
Chidambaram et al., 2004	Public Health	Ethiopia	Mass azithromycin distribution	Trachoma	Unmatched cohort	Within-cluster spillover effect (4)
Chong et al., 2013	Economics	Colombia	Online sexual health education	Knowledge about sexually transmitted infections	Double-randomized trial	Total effect (6), Within-cluster spillover effect (4), Direct effect conditional on treatment density (18a)
Contreras & Maitra, 2013	Economics	Colombia	Conditional cash transfers	Self-reported illness	Cohort study + propensity-score matching + difference-in-differences	Within-cluster spillover effect among ineligibles (8)
Cooper & Fitch, 1983	Public Health	St. Lucia	Pertussis vaccine	Pertussis	Cross-sectional survey	Treatment coverage mean (1)
Dupas, 2006	Economics	Kenya	Information about HIV transmission	Teen pregnancy	Re-analysis of a cluster-randomized trial	Total effect (6), Total effect conditional on treatment density (21)
Egere et al., 2012	Public Health	The Gambia	Pneumococcal conjugate vaccine	Pneumococcal nasopharyngeal carriage	Cluster-randomized trial	Within-cluster spillover effect among ineligibles (8)
Emch et al., 2006	Geography	Bangladesh	Cholera vaccine	Cholera	Re-analysis of an individually randomized trial	Treatment coverage effect (2)
Emch et al., 2009	Geography	Bangladesh	Cholera vaccine	Cholera	Re-analysis of an individually randomized trial	Treatment coverage mean (1), Treatment coverage effect (2)
Fitzsimons et al., 2012	Economics	Malawi	Information on infant nutrition and health	Child growth and morbidity	Re-analysis of a cluster-randomized trial	Total effect among eligibles (10), Within-cluster spillover effect among ineligibles (8)

Forleo-Neto et al., 1999	Public Health	Brazil	*Haemophilus influenzae* type b vaccine	*Haemophilus influenzae* type b carriage	Cross-sectional survey	Treatment coverage mean (1)
German et al., 2012	Public Health	Thailand	Peer network health education	Depression	Individually randomized trial	Direct effect (3), Spillover effect among social network members (26)
Godlonton & Thornton, 2012	Economics	Malawi	Incentives for voluntary counseling and testing for HIV	Voluntary counseling and testing for HIV	Re-analysis of an individually randomized trial + instrumental variables	Total effect conditional on outcome density (22)
Haile et al., 2013	Public Health	Ethiopia	Latrines	Trachoma	Re-analysis of a cluster-randomized trial	Treatment coverage effect (2)
Hammitt et al., 2014	Public Health	Kenya	Pneumococcal conjugate vaccine	Pneumococcal nasopharyngeal carriage	Cross-sectional surveys before and after intervention	Spillover before and after treatment (29)
Handa et al., 2001	Economics	Mexico	Conditional cash transfers	Child nutrition surveillance	Re-analysis of a cluster-randomized trial	Total effect among eligibles (10), Within-cluster spillover effect among ineligibles (8)
Hawley et al., 2003	Public Health	Kenya	Insecticide-treated nets	Child mortality, anemia, clinical malaria	Re-analysis of a cluster-randomized trial	Spillover effect conditional on distance to nearest treated cluster (14), Direct effect conditional on distance to nearest treated cluster (15)
House et al., 2009	Public Health	Ethiopia	Mass azithromycin distribution	Trachoma	Cluster-randomized trial	Within-cluster spillover effect among ineligibles (8)
Huq et al., 2010	Public Health	Bangladesh	Water filtration with sari cloth and nylon cloth	Cholera	Cross-sectional survey of a population that previously participated in an individually randomized trial	Treatment coverage mean (1), Direct effect (3)
Janssens et al., 2005	Economics	India	Women's empowerment program	Vaccine coverage	Cohort study + propensity score matching + instrumental variables	Within-cluster spillover effect (4), Direct effect (3), Within-cluster spillover effect in which controls are matched to the untreated (4)
Joshi & Shultz, 2013	Economics	Bangladesh	Maternal and child health program	Parity, body mass index activities daily index, childhood vaccination	Cross-sectional survey of a population previously in an unmatched cohort study	Spillover effect into boundary areas of untreated clusters (17)

Kazianga et al., 2014	Economics	Burkina Faso	School feeding program	Child growth	Cluster-randomized trial	Total effect among eligibles (10), Within-cluster spillover effect among ineligibles (8)
Khan et al., 2012	Public Health	Pakistan	Typhoid vaccine	Typhoid fever	Cluster-randomized trial	Direct effect (3), Within-cluster spillover effect (4), Total effect (6)
Khatib et al., 2012	Public Health	Tanzania	Cholera vaccine	Cholera	Cohort study without a control group	Direct effect (3), Treatment coverage mean (1)
Kremer & Miguel, 2007	Economics	Kenya	Exposure to information about deworming	Deworming	Cross-sectional survey of a population that previously participated in a cluster-randomized trial	Total effect conditional on number of social network links (27)
Miguel & Kremer, 2004	Economics	Kenya	Deworming	Helminth infection	Cluster-randomized trial	Total effect (6), Within-cluster spillover effect (4), Spillover effect conditional on treatment density (19b)
Ozier, 2011	Economics	Kenya	Deworming	Child growth and cognitive performance	Cross-sectional survey of a population that previously participated in a cluster-randomized trial	Within-cluster spillover effect (4)
Paul et al., 1962	Public Health	Costa Rica	Polio vaccine	Polio antibody conversion	Individually randomized trial	Vaccine efficacy** (18)
Perez-Heydrich et al., 2014	Public Health	Bangladesh	Cholera vaccine	Cholera	Re-analysis of a cluster-randomized trial + inverse probability weighting	Direct effect conditional on treatment density 18), Spillover effect conditional on treatment density (19), Total effect conditional on treatment density (21)
Préziosi & Halloran, 2003	Public Health	Senegal	Pertussis vaccine	Pertussis	Case-control study	Vaccine efficacy for infectiousness (24), Vaccine efficacy for susceptibility (23), Total vaccine efficacy (25)
Ribas et al., 2011	Economics	Paraguay	Conditional cash transfers	Health care visits	Cohort study + propensity-score matching + difference-in-differences	Direct effect (3), Within-cluster spillover effect among ineligibles (8)
Roca et al., 2011	Public Health	The Gambia	Pneumococcal conjugate vaccine	Pneumococcal nasopharyngeal carriage	Cluster-randomized trial	Total effect conditional on exposure to treatment (11)
Roca et al., 2013	Public Health	The Gambia	Pneumococcal conjugate vaccine	Pneumococcal nasopharyngeal carriage	Cluster-randomized trial	Total effect conditional on exposure to treatment (11)

Root et al., 2011	Geography	Bangladesh	Cholera vaccine	Cholera	Re-analysis of an individually randomized trial	Treatment coverage effect (2)
Root et al., 2014	Geography	Philippines	Pneumococcal conjugate vaccine	Pneumonia	Re-analysis of a cluster-randomized trial	Treatment coverage effect (2)
Shekhawat et al., 2014	Public Health	Tanzania	Mass azithromycin distribution	Trachoma	Cohort study in a population previously enrolled in a cluster-randomized trial	Within-cluster spillover effect** (4)
Singh (year not listed)	Economics	India	Nutrition education	Child growth	Cluster-randomized trial	Total effect (6), Within-cluster spillover effect (4)
Sur et al., 2009	Public Health	India	Typhoid vaccine	Typhoid fever	Cluster-randomized trial	Total effect (6), Within-cluster spillover effect (4)
Tontarawongsa et al., 2011	Economics	India	Insecticide-treated nets for free or with microloans and information sessions	Insecticide-treated net use	Re-analysis of a cluster-randomized trial	Total effect conditional on number of social network links (27), Spillover effect among social network members (26)
Ziegelhöfer, 2012	Economics	Guinea	Improved water supply	Diarrhea	Cross-sectional survey + regression discontinuity + instrumental variables	Total effect (6), Spillover effect conditional on treatment density (19), Ratio of village level to household level effect (28)
Zivin et al., 2009	Economics	Kenya	HIV/AIDS treatment	Child growth	Cohort study	Within-cluster spillover effect (4)

*If multiple types of primary outcomes are listed, the primary health outcomes are mentioned here.
**This parameter was not explicitly estimated, but it could have been using the data collected in the study.

Table 4: Study designs by academic discipline

Design	Economics	Geography	Public Health	Total
Cluster-randomized trial	5	0	8	13
Individually randomized trial	0	0	2	2
Double-randomized trial	2	0	0	2
Re-analysis of a cluster-randomized trial	5	1	2	8
Re-analysis of a cluster-randomized trial + inverse probability weighting	0	0	1	1
Re-analysis of an individually randomized trial	1	3	2	6
Re-analysis of an individually randomized trial + instrumental variables	1	0	0	1
Cohort study in a population previously enrolled in a cluster-randomized trial	0	0	1	1
Cross-sectional survey of a population that previously participated in a cluster-randomized trial	3	0	0	3
Cross-sectional survey of a population that previously participated in an individually randomized trial	0	0	1	1
Cross-sectional surveys before and after intervention	0	0	1	1
Cohort study	1	0	0	1
Cohort study + propensity-score matching + difference-in-differences	3	0	0	3
Cohort study + propensity score matching + instrumental variables	1	0	0	1
Cross-sectional survey + regression discontinuity + instrumental variables	1	0	0	1
Cross-sectional survey of a population previously in an unmatched cohort study	2	0	0	2
Case-control study	0	0	2	2
Re-analysis of matched case-control study	0	0	1	1
Cohort study without a control group	0	0	1	1
Cross-sectional survey	0	0	2	2
Unmatched cohort	0	0	1	1
Total	25	4	25	54

Table 5: Intervention types by academic discipline

	Economics	Geography	Public Health	Total
Cholera vaccine	0	3	5	8
Community monitoring and provision of health services	1	0	0	1
Conditional cash transfers	5	0	0	5
Deworming	2	0	0	2
Exposure to information about deworming	1	0	0	1
Haemophilus influenzae type b vaccine	0	0	2	2
HIV/AIDS treatment	1	0	0	1
Improved water supply	1	0	0	1
Incentives for voluntary counseling and testing for HIV	1	0	0	1
Information about HIV transmission	1	0	0	1
Information on infant nutrition and health	1	0	0	1
Insecticide-treated nets	0	0	1	1
Insecticide-treated nets for free or with microloans and information sessions	1	0	0	1
Latrines	0	0	1	1
Mass azithromycin distribution	0	0	3	3
Maternal and child health program	2	0	0	2
Nutrition education	1	0	0	1
Online sexual health education	1	0	0	1
Peer network health education	0	0	1	1
Pertussis vaccine	0	0	3	3
Pneumococcal conjugate vaccine	0	1	4	5
Polio vaccine	0	0	1	1
School feeding program	2	0	0	2
Subsidized deworming	1	0	0	1
Subsidized insecticide-treated nets	1	0	0	1
Typhoid vaccine	0	0	2	2
Vaccines (BCG, DPT, polio, measles) + incentives	1	0	0	1
Water filtration with sari cloth and nylon cloth	0	0	1	1
Women's empowerment program	1	0	0	1
Women's groups and health service strengthening	0	0	1	1
Total	25	4	25	54

Table 6: Intervention types by reason for cluster vs. individual level treatment allocation

	No rationale given	Evaluation allocation scheme based on level of intervention delivery	Treatment was not allocated by the investigators	Treatment was allocated to clusters due to political and logistical reasons	Spillover measurement was not the primary reason, but it was included in the design phase	Treatment was allocated to clusters to be consistent with past trials for this intervention	Clusters were randomized in phases as funds became available	Treatment was allocated to clusters in order to measure spillovers	Total
Conditional and unconditional cash transfers	0	0	0	0	0	0	0	1	1
Conditional cash transfers	1	3	0	0	0	0	0	0	4
Health education	0	2	0	0	0	0	1	2	5
Health systems	0	1	0	0	0	0	0	0	1
HIV/AIDS	0	1	1	0	0	0	0	0	2
Insecticide-treated nets	1	0	0	0	0	2	0	0	3
Mass drug administration (for parasite control)	1	6	0	0	0	0	0	0	7
Maternal and child health	2	0	0	0	0	0	0	0	2
Nutrition	0	1	0	1	0	0	0	0	2
Peer support intervention	0	0	0	0	1	0	0	0	1
Vaccines	9	0	6	0	2	0	0	5	22
Water and sanitation	2	0	1	0	0	0	0	0	3
Women's empowerment program	0	1	0	0	0	0	0	0	1
Total	16	15	8	1	3	2	1	8	54

Table 7: Spillover mechanism by academic discipline

	Economics (n=25)	**Geography (n=4)**	**Public Health (n=21)**	**All (n=50)**
Geographic proximity	48%	75%	96%	72%
Social proximity	56%	25%	8%	31%
Learning/imitation	52%	0%	8%	28%
Norm-shaping	52%	0%	8%	28%
Income/substitution effect	8%	0%	0%	4%
Public good effect	28%	0%	0%	13%
General equilibrium effects	4%	0%	0%	2%
Relative deprivation	4%	0%	0%	2%

Table 8: Spillover parameters estimated in included studies and related direct and total effects parameters

	Parameter	Definition	Notes
	Treatment coverage parameters		
1	Treatment coverage mean	$E[Y_c \mid P_c = p_c]$ for $c = 1,...,C$	Y_c = the mean outcome in area c P_c = the proportion treated in area c
2	Treatment coverage effect	$E[Y_c \mid P_c = p_c]$-$E[Y_c \mid P_c = p_c - \delta]$ for $c = 1,...,C$	δ = a pre-determined difference in p_c
	Within-cluster spillovers and related parameters		
3	Direct effect	$E[Y \mid X=1, T=1]$ - $E[Y \mid X=0, T=0]$	Y = the outcome
4	Within-cluster spillover effect*	$E[Y \mid X=1, T=0]$ - $E[Y \mid X=0, T=0]$	
5	Within-treated-cluster spillover effect	$E[Y \mid X=1, T=1]$ - $E[Y \mid X=1, T=0]$	T = indicator for whether an individual was allocated to treatment
6	Total effect	$E[Y \mid X=1]$ - $E[Y \mid X=0]$	
7	Direct effect among eligibles	$E[Y \mid X=1, T=1, E=1]$ - $E[Y \mid X=0, T=0, E=1]$	X = indicator for whether a cluster was treated
8	Within-cluster spillover effect among ineligibles	$E[Y \mid X=1, T=0, E=0]$ - $E[Y \mid X=0, T=0, E=0]$	
9	Within-treated-cluster spillover effect among ineligibles	$E[Y \mid X=1, T=1, E=0]$ - $E[Y \mid X=1, T=0, E=0]$	E = indicator for whether an individual was eligible for treatment
10	Total effect among eligibles	$E[Y \mid X=1, E=1]$ - $E[Y \mid X=0, E=1]$	M = indicator for whether an individual was exposed to treatment
11	Total effect conditional on exposure to treatment	$E[Y \mid X=1, T=1, M=1]$ - $E[Y \mid X=0, T=0]$	
12	Within-cluster spillover effect conditional on exposure to treatment	$E[Y \mid X=1, T=0, M=1]$ - $E[Y \mid X=0, T=0]$	
	Distance-based parameters		
13	Spillover effect conditional on living in an untreated cluster within distance d to treated clusters	$E[Y \mid X=0, D_t=1]$-$E[Y \mid X=0, D_t=0]$	D_t = indicator for whether an individual lived in within distance d to the nearest treated cluster
14	Spillover effect conditional on distance to nearest treated cluster	$E[Y \mid X=0, T=0, D_t=1]$-$E[Y \mid X=0, T=0, D_c=1]$	D_c = indicator for whether an individual lived within distance d to the nearest control cluster
15	Direct effect conditional on distance to nearest control cluster	$E[Y \mid X=1, T=1, D_c=1]$-$E[Y \mid X=0, T=0, D_c=1]$	D_{tc} = indicator for whether treatment and control units were within a distance d from each other
16	Total effect conditional on whether treatment and control units were within distance d of	$E[Y \mid X=1, D_{tc}=1]$-$E[Y \mid X=0, D_{tc}=0]$	B = indicator for living in the boundary area of an

	each other		untreated cluster close to a treated cluster
17	Spillover effect into boundary areas of untreated clusters	E[Y |X=0, B=1]-E[Y |X=0, B=0]	

	Parameter	Definition	
	Parameters conditional on density		
18a	Direct effect conditional on treatment density (a)	E[Y |X=1, T=1, $P(d)$=$p(d)$+ δ] – E[Y |X=0, T=0, $P(d)$=$p(d)$]	$P(d)$ = proportion treated within distance d
18b	Direct effect conditional on treatment density (b)	E[Y |X=1,T=1, $N_t(d)$=$n_t(d)$+δ, $N(d)$=$n(d)$ +δ] - E[Y |X=0,T=0, $N_t(d)$=$n_t(d)$, $N(d)$=$n(d)$]	$N_t(d)$ = number of people treated within dist d
19a	Spillover effect conditional on treatment density (a)	E[Y |X=0, $P(d)$=$p(d)$+ δ] – E[Y |X=0, $P(d)$=$p(d)$]	$N(d)$ = number of people within distance d
19b	Spillover effect conditional on treatment density (b)	E[Y |X=0, $N_t(d)$=$n_t(d)$+δ, $N(d)$=$n(d)$ +δ] – E[Y |X=0, $N_t(d)$=$n_t(d)$, $N(d)$=$n(d)$]	Px = indicator for treatment density
20	Within-treated-cluster spillover effect conditional on treatment density	E[Y |X=1, T=1, Px=px + δ] – E[Y |X=1, T=0, Px=px]	Py = indicator for outcome density
21	Total effect conditional on treatment density	E[Y |X=1, Px=px + δ]-E[Y |X=0, Px=px]	δ = a pre-determined difference in px
22	Total effect conditional on outcome density	E[Y |X=0, Py=py + δ]-E[Y |X=0, Py=py]	
	Spillovers conditional on exposure to cases		
23	Vaccine efficacy for susceptibility (VE_S)	1 - [(E[Y_i |T_i=1,Y_j =1]) / (E[Y_i |T_i=0,Y_j =1])]	Y_i = outcome of susceptible individuals Y_j = outcome among cases
24	Vaccine efficacy for infectiousness (VE_I)	1 - [(E[Y_i |T_j=1,Y_j =1]) /	T_i = treatment among susceptible individual

		$(E[Y_i \mid T_j=0, Y_j=1])]$	T_j = treatment among cases
25	Total vaccine efficacy	$1 - (E[Y_i \mid T_j=1, T_j=1, Y_j=1] /$ $E[Y_i \mid T_j=1, T_j=1, Y_j=1])$	
	Social network spillover parameters		
26	Spillover effect among social network members	$E[Y_i \mid T_i=0, T_j=1] - E[Y_i \mid T_i=0, T_j=0]$	Y_i = outcome of alters Y_j = outcome of egos
27	Total effect conditional on number of social network links	$E[Y \mid X=1, N=n+\delta] - E[Y \mid X=0, N=n]$	T_i = treatment of alters T_j = treatment of egos N = number of social network links δ = a pre-determined difference in N
	Other spillover parameters		
28	Ratio of village level to household level effects	$E[Y \mid X=1] - E[Y \mid X=0] /$ $E[Y \mid T=0, X=1] - E[Y \mid X=0, T=0,]$	T = indicator for whether an individual was allocated to treatment X = indicator for whether a cluster was treat
29	Spillover before and after treatment	$E[Y \mid T=0, Z=1] - E[Y \mid T=0, Z=0]$	Z = indicator for time of treatment

* The parameter for within-cluster spillovers that matches individuals in the control group to those who were untreated in the treatment group can be expressed with this notation as well.

Table 9: Studies estimating treatment-density spillovers

Paper	Intervention	Health-Related Outcomes	Design	Density measure
Effects conditional on treatment density				
Baird et al., 2013*	Subsidized deworming	Self-reported health and BMI	Cross-sectional survey of a population that previously participated in a cluster-randomized trial	Proportion of treated students within 6km of a school
Bhattacharya et al., 2013	Subsidized insecticide-treated nets	Purchase of insecticide-treated nets	Re-analysis of a randomized trial	Proportion of treated households residing in 250m, 500m, and 1000m radius
Chong et al., 2013	Online sexual health education	Sexually transmitted infections (STIs), knowledge of STIs, redemption of condom vouchers	Double-randomized trial	Proportion of students' closest friends in study schools who were treated
Dupas, 2006	Information about HIV transmission	Teen pregnancy	Re-analysis of a cluster-randomized trial	Proportion of treated classmates in the study of the same sex
Perez-Heydrich et al., 2014	Cholera vaccine	Cholera	Re-analysis of a cluster-randomized trial + inverse probability weighting	Neighborhoods defined using a single linkage agglomerative clustering method
Miguel & Kremer, 2004*	School-based deworming	Soil-transmitted helminth infection	Cluster-randomized trial	Proportion of treated students within 6km of a school
Ziegelhöfer, 2012	Improved water supply	Diarrhea	Cross-sectional survey + regression discontinuity + instrumental variables	Fraction of individuals receiving treatment within a geographic area (area size not stated)
Effects conditional on outcome density				
Godlonton & Thornton	Incentives for voluntary counseling and testing for HIV	Voluntary counseling and testing for HIV	Re-analysis of a randomized trial + instrumental variables	Proportion of neighbors within 0.5 km who were tested for HIV

*Studies conducted using the same primary dataset

Table 10: Search terms related to spillover effects in included texts by academic field†

	Economics	Geography	Public Health	Total
Indirect effect*	12	2	13	27
Spillover*	23	0	1	24
Externalit*	19	0	0	19
Seconda*	3	3	10	16
Indirect protection	0	4	11	15
Herd protect*	0	2	12	14
Diffusion	7	1	3	11
Herd immunity	1	4	5	10
Herd effect*	0	0	10	10
Peer effect*	9	0	0	9
Unexpected	2	0	3	5
Interference	2	0	2	4
Indirect protective	0	0	4	4
Contagion	3	0	0	3
Unexpected benefit*	0	0	1	1

* Asterisks at the end of search terms indicate wild card characters allowed at the end of the search term. For example, “externalit*” would retrieve search results for “externality” and “externalities”.

† Counts allow for multiple terms per included text

Figure 1: Double-randomized design parameters

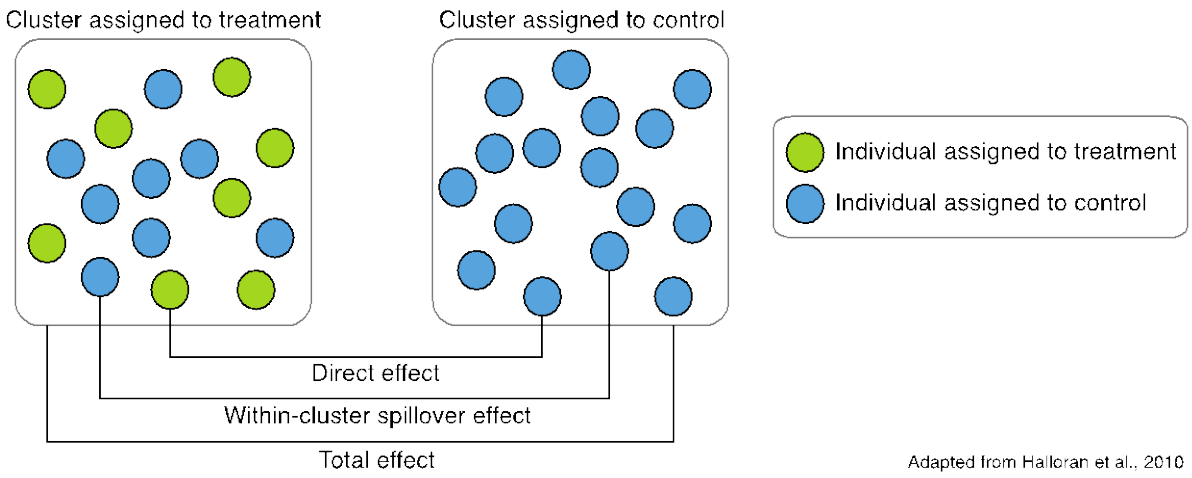


Figure 2: Within-cluster spillover effects

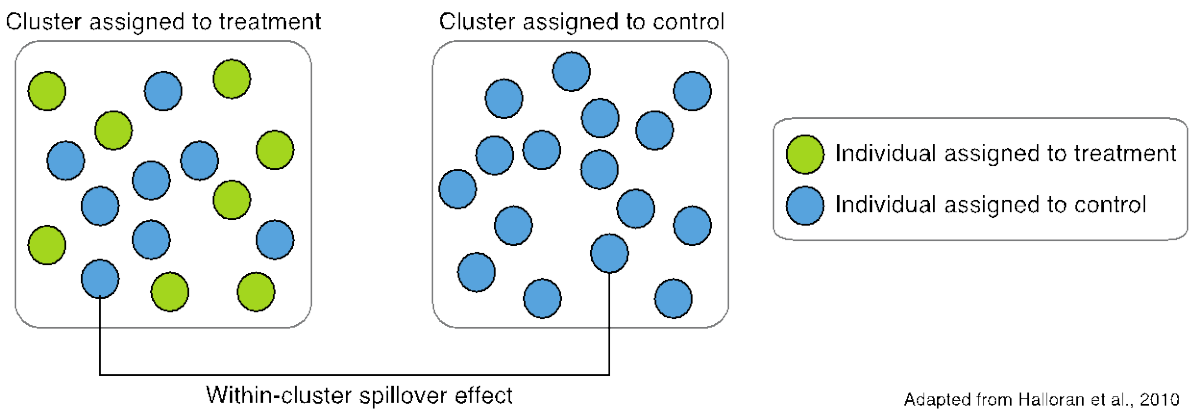


Figure 3: Distanced-based spillover effect

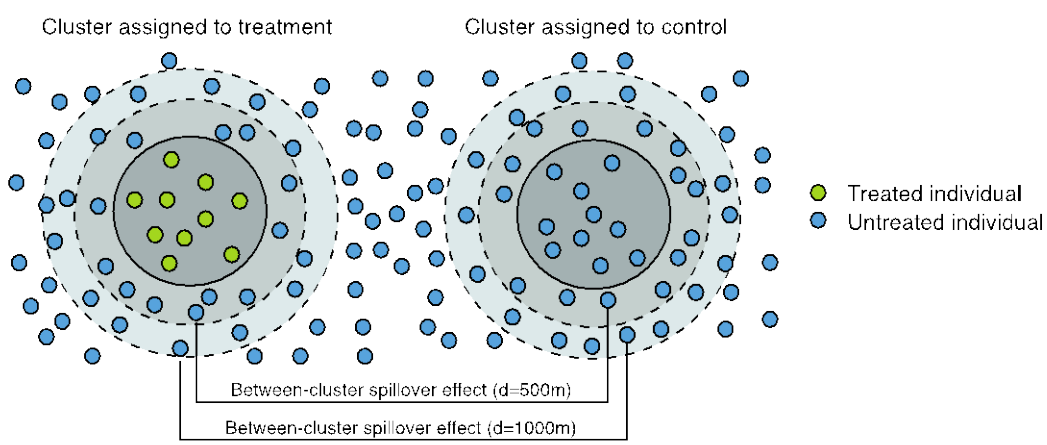


Figure 4: Spillovers conditional on treatment density

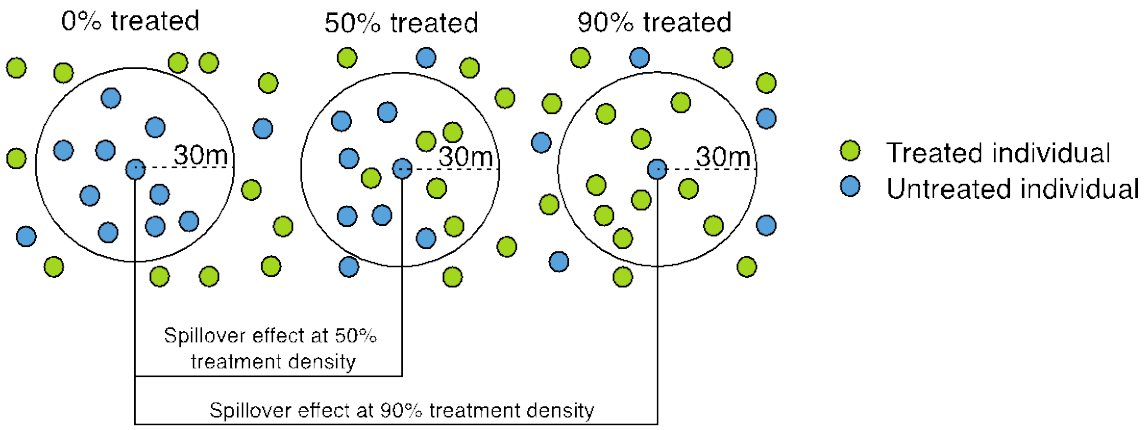


Figure 5: Spillovers conditional on exposure to cases

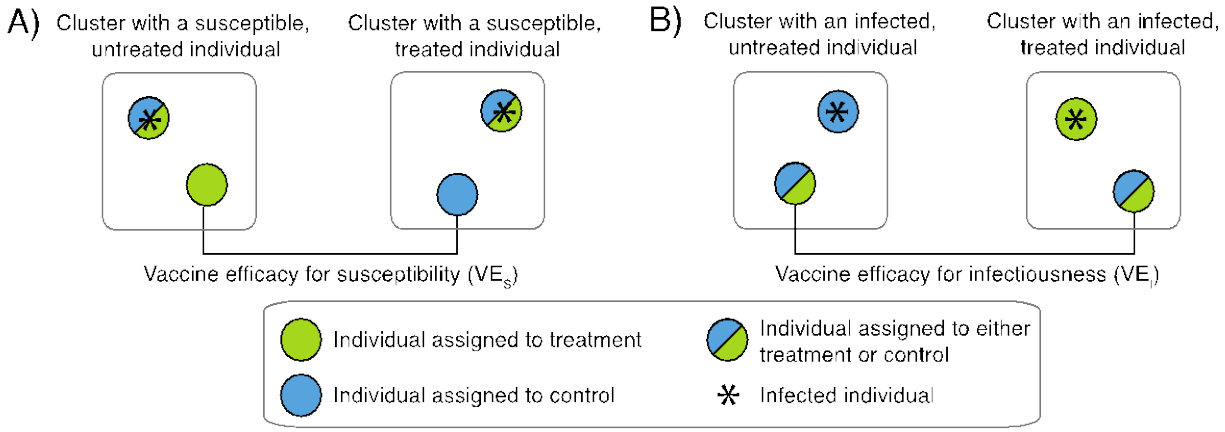


Figure 6: Social network spillovers

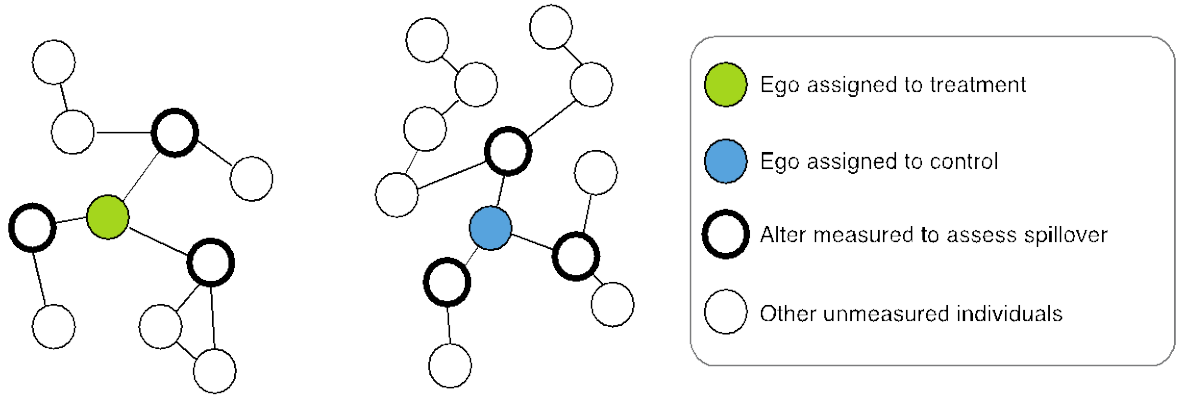


Figure 7: Records in each stage of the systematic review

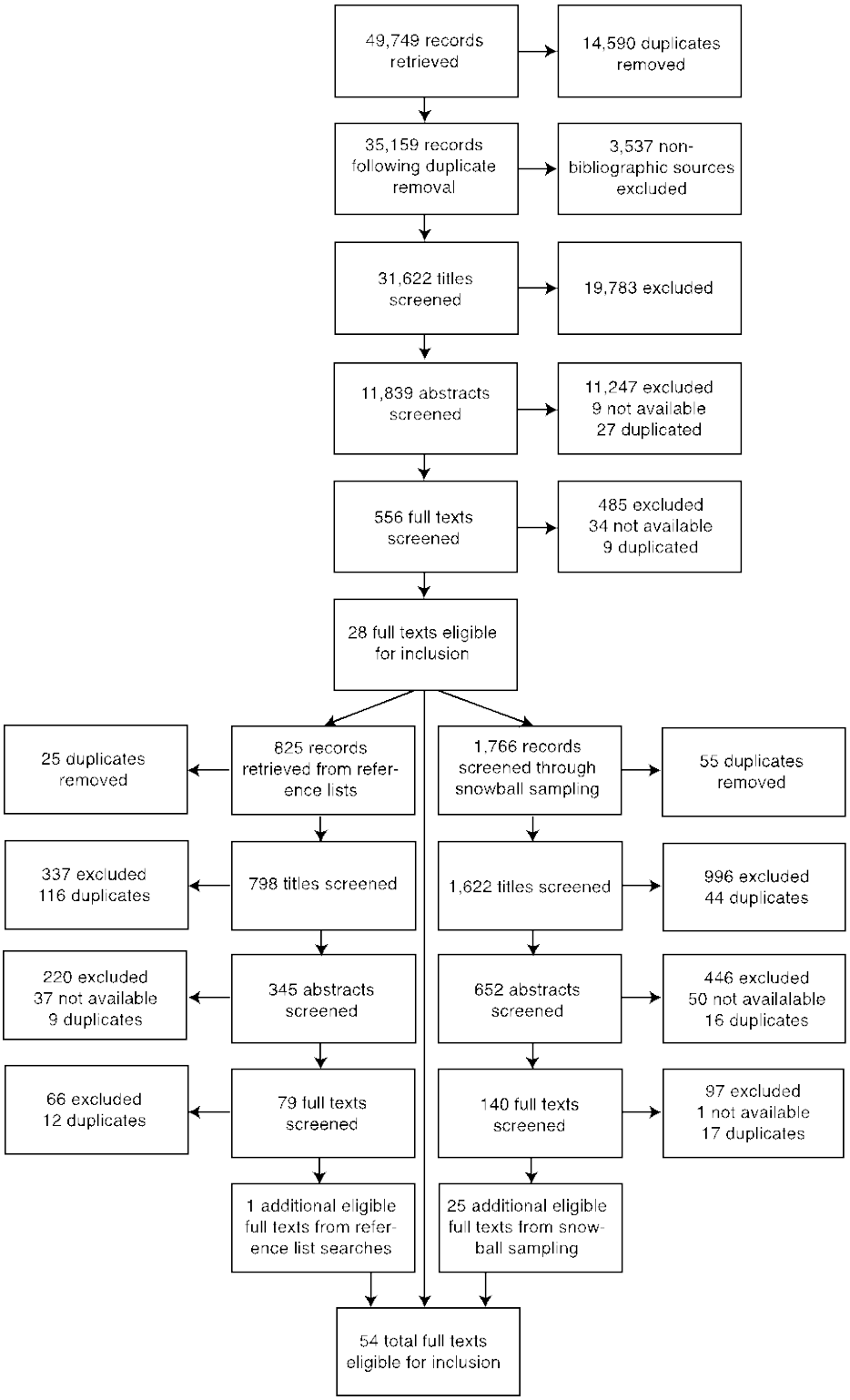
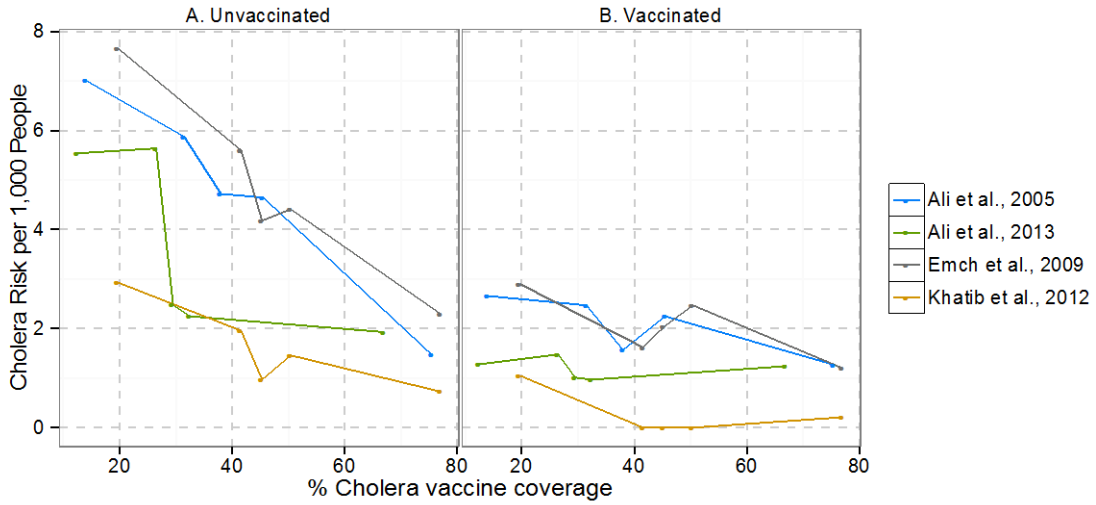


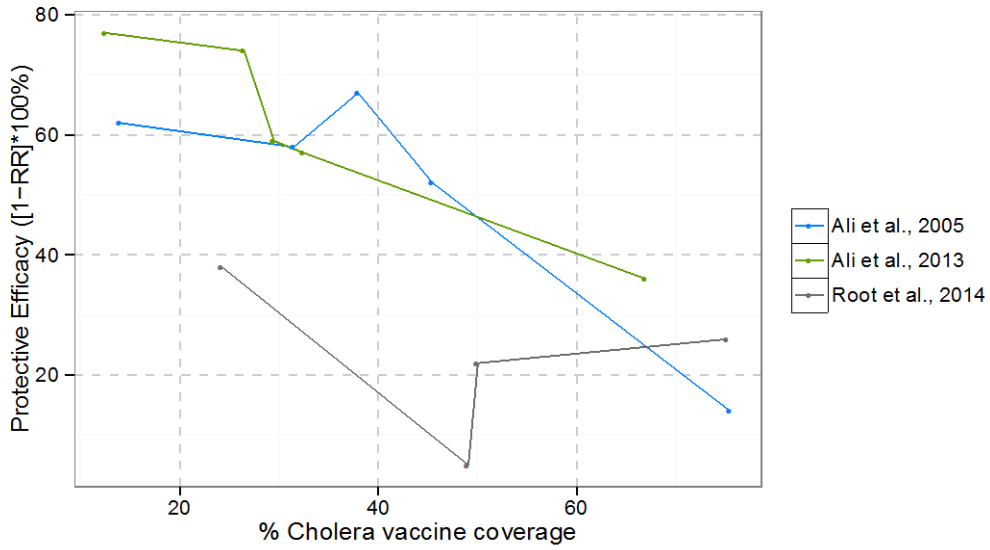
Figure 8: Number of spillover parameters estimated by discipline

Figure 9: Cholera risk per 1,000 people among unvaccinated and vaccinated individuals by varying levels of cholera vaccine coverage



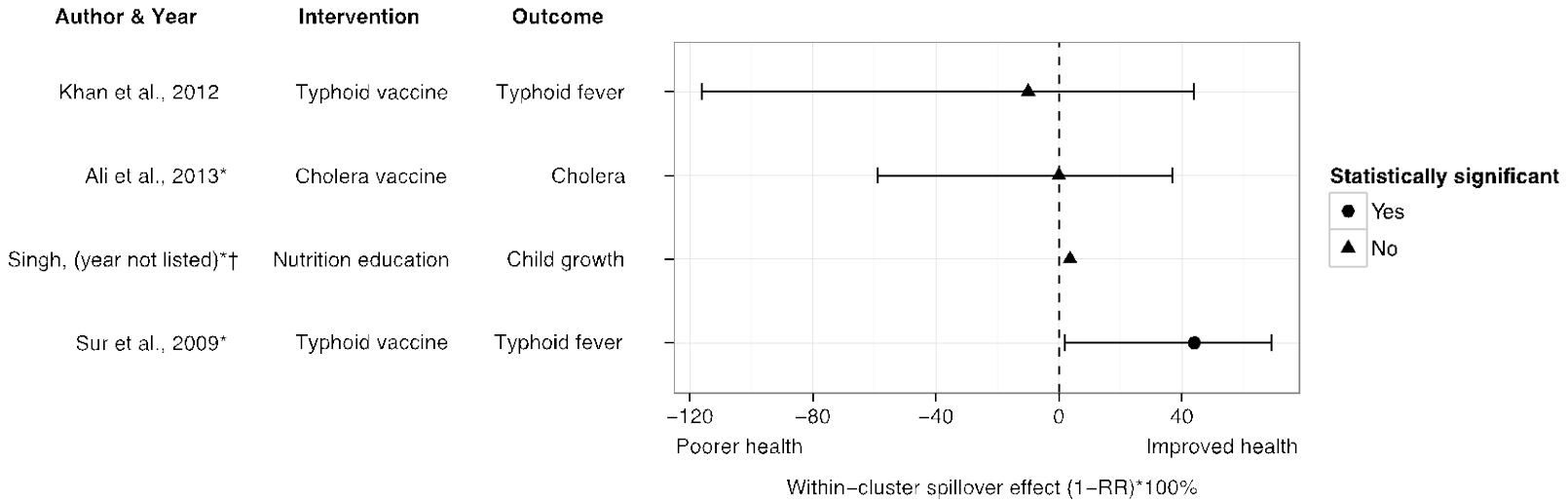
Cholera risk decreases at higher levels of vaccine coverage among the unvaccinated and stays relatively the same across levels of vaccine coverage among the vaccinated. These patterns indicate herd protection; unvaccinated individuals living in areas with higher vaccine coverage are protected from infection due to reduced transmission. Cholera risk is similar among vaccinated individuals across levels of cholera vaccine coverage because vaccinated individuals are directly protected from infection from the vaccine they received.

Figure 10: Cholera vaccine protective efficacy by varying levels of cholera vaccine coverage



The protective efficacy decreases at higher levels of vaccine coverage because as vaccine coverage increases, the risk among unvaccinated decreases, as shown in Figure 8; thus, the difference in risk between the vaccinated and unvaccinated individuals decreases as vaccination coverage increases, and thus the relative risk is closer to the null and the protective efficacy is closer to zero.

Figure 11: Within-cluster spillovers

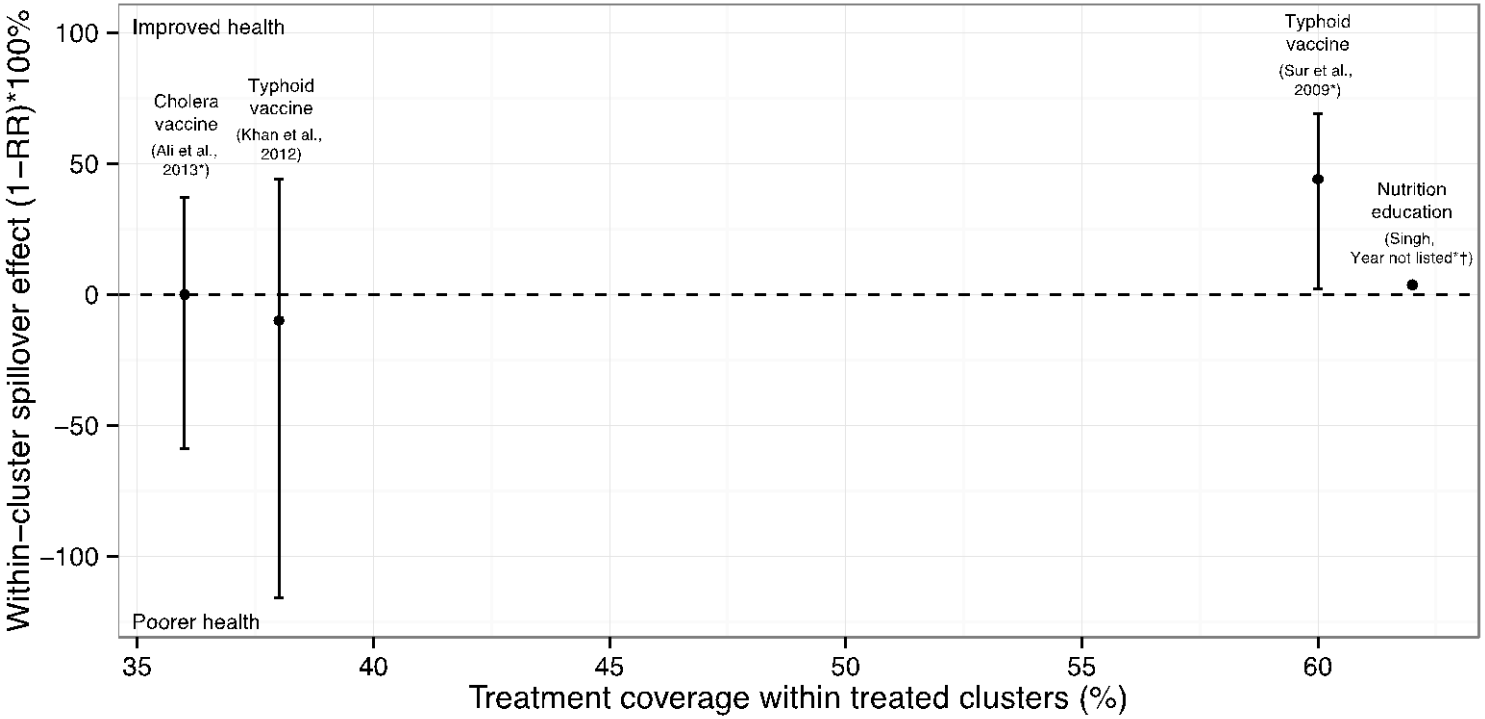


On the x-axis, the within-cluster spillover effect is shown as the percent change in outcome among the untreated in the treated cluster from the mean in the control group. Outcomes were recoded so that a greater value of the spillover effect indicates an improvement in health (e.g. more vaccination, lower mortality) and a smaller value indicates poorer health (e.g. less vaccination, higher mortality). This figure excludes studies with overall high risk of bias or which did not report information that allowed for standardization. Statistical significance was determined based on the measures presented in the paper for the parameter on its original scale.

* These studies were conducted in the same country (India) and are subject to dependence.

† Information required to convert standard errors for risk differences to standard errors for (1-RR)*100% was not reported, thus 95% confidence intervals are not presented.

Figure 12: Within-cluster spillovers by level of treatment coverage

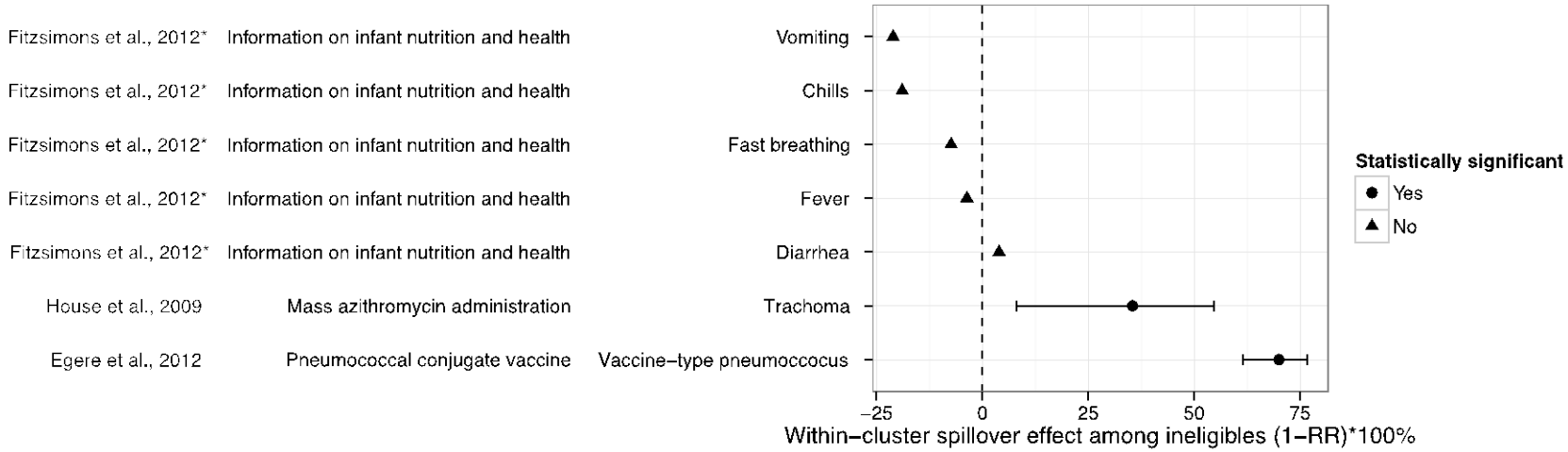


On the y-axis, the within-cluster spillover effect is shown as the percent change in outcome among the untreated in the treated cluster from the mean in the control group. Outcomes were recoded so that a greater value of the spillover effect indicates an improvement in health (e.g. more vaccination, lower mortality) and a smaller value indicates worse health (e.g. less vaccination, higher mortality). This figure excludes studies with overall high risk of bias or which did not report information that allowed for standardization.

* These studies were conducted in the same country (India) and are subject to dependence.

† Information required to convert standard errors for risk differences to standard errors for (1-RR)*100% was not reported, thus 95% confidence intervals are not presented.

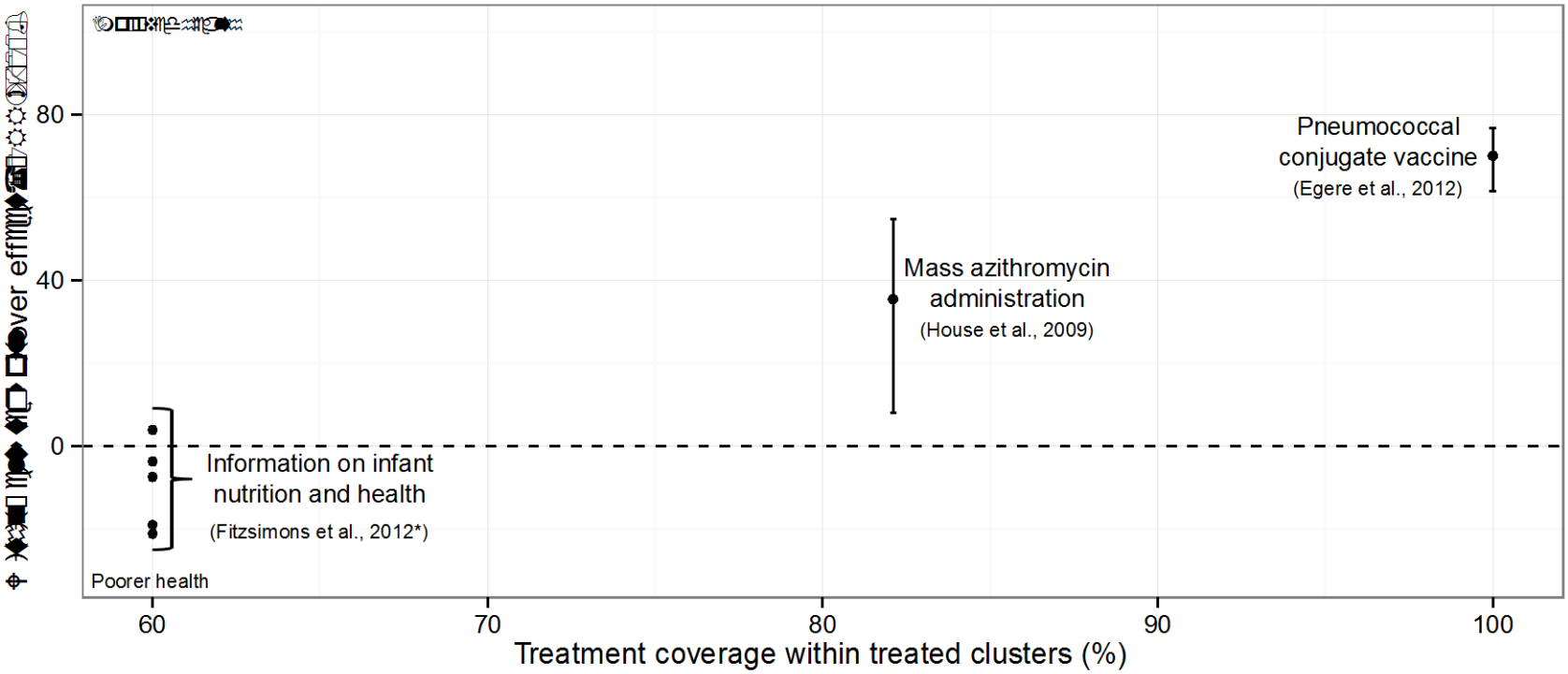
Figure 13: Within-cluster spillovers among individuals who were ineligible to receive the intervention



On the x-axis, the within-cluster spillover effect is shown as the percent change in outcome among the untreated in the treated cluster from the mean in the control group. Outcomes were recoded so that a greater value of the spillover effect indicates an improvement in health (e.g. more vaccination, lower mortality) and a smaller value indicates poorer health (e.g. less vaccination, higher mortality). This figure excludes studies with overall high risk of bias or which did not report information that allowed for standardization. Statistical significance was determined based on the measures presented in the paper for the parameter on its original scale.

* Information required to convert standard errors for risk differences to standard errors for (1-RR)*100% was not reported, thus 95% confidence intervals are not presented.

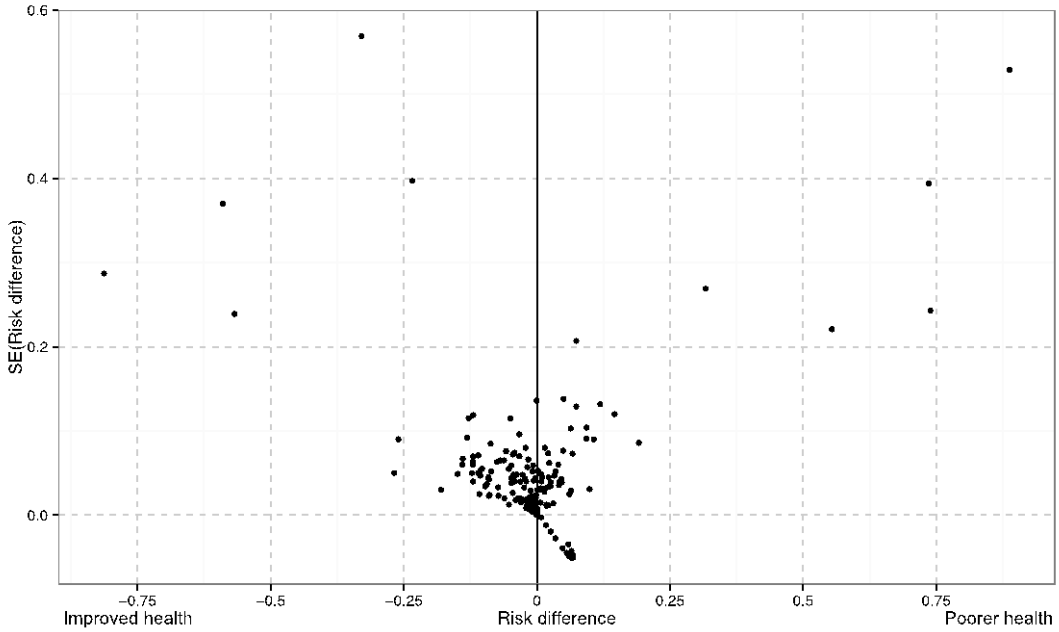
Figure 14: Within-cluster spillovers among individuals who were ineligible to receive the intervention by level of treatment coverage



On the y-axis, the within-cluster spillover effect among ineligibles is shown as the percent change in outcome among the untreated in the treated cluster from the mean in the control group. Outcomes were recoded so that a greater value of the spillover effect indicates an improvement in health (e.g. more vaccination, lower mortality) and a smaller value indicates worse health (e.g. less vaccination, higher mortality). This figure excludes studies with overall high risk of bias or which did not report information that allowed for standardization.

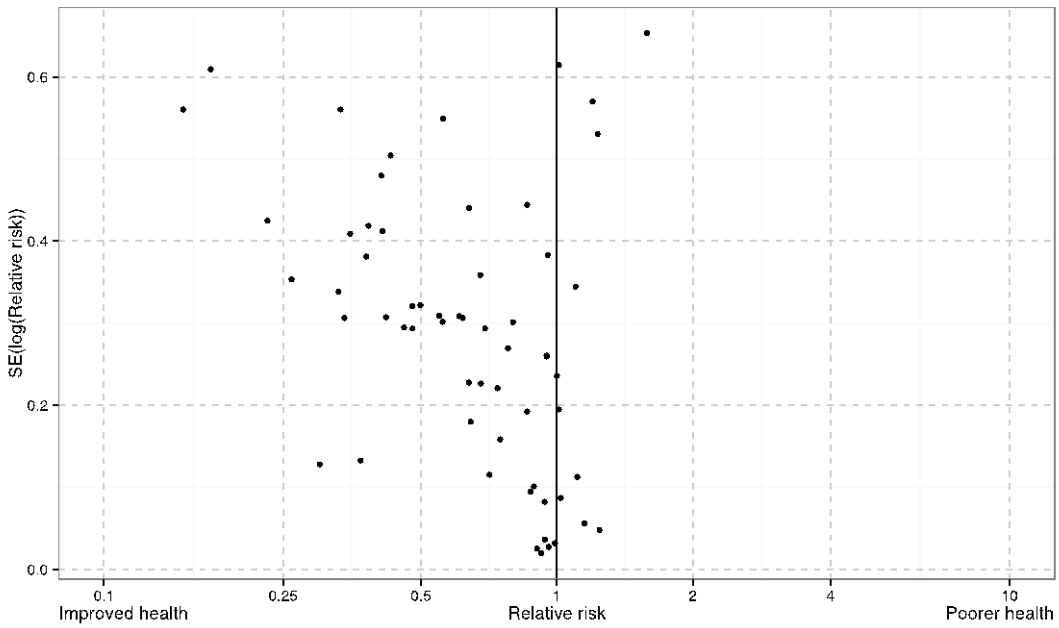
*Information required to convert standard errors for risk differences to standard errors for (1-RR)*100% was not reported, thus 95% confidence intervals are not presented.

Figure 15a: Funnel plot for spillover effects for studies estimating risk differences



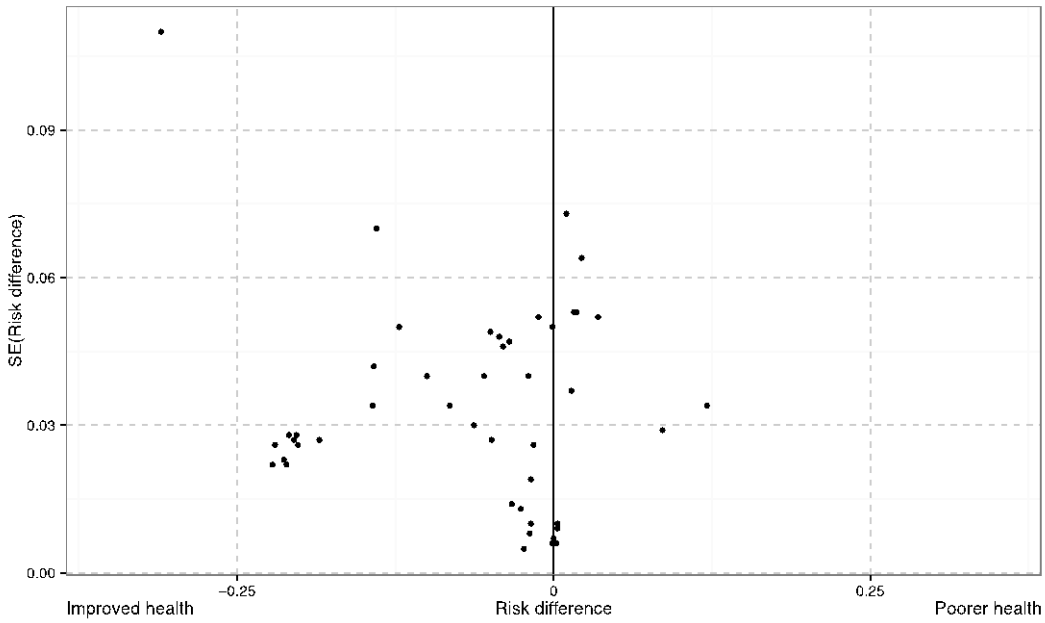
This plot includes spillover estimates from 18 studies that reported risk differences for binary outcomes, of which all but one were in the economics literature. These studies evaluated a wide range of interventions including women's empowerment programs, mass drug administration for parasite control, peer group interventions, and nutrition programs.

Figure 15b: Funnel plot for spillover effects for studies estimating risk ratios



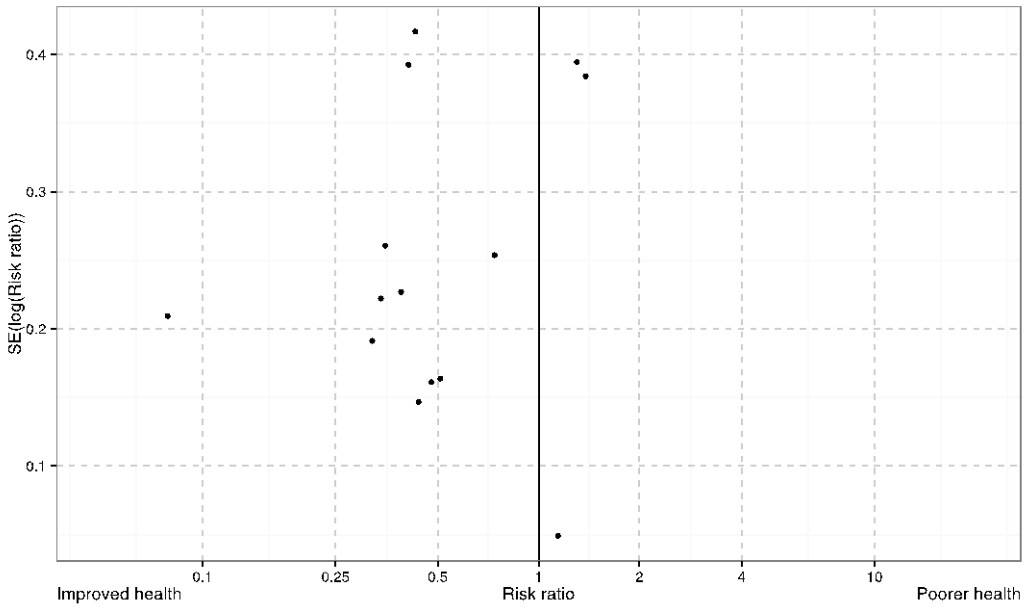
This plot includes spillover estimates from 10 studies that reported risk ratios or protective efficacy ((1-RR)x100%) for binary outcomes, all of which were in the public health literature. These studies evaluated a vaccines and mass drug administration for parasite control.

Figure 15 c: Funnel plot for direct and total effects for studies estimating risk differences



This plot includes total and direct effect estimates from 11 studies that reported risk differences for binary outcomes, of which all but one were in the economics literature. These studies evaluated a wide range of interventions including women's empowerment programs, mass drug administration for parasite control, peer group interventions, and nutrition programs.

Figure 15 d: Funnel plot for direct and total effects for studies estimating risk ratios



This plot includes total and direct effect estimates from 5 studies that reported risk ratios for binary outcomes, all of which were in the public health literature. These studies evaluated a vaccines and mass drug administration for parasite control.

Appendix 1: Search strategy

Bibliographic search strategy

CAB/Ovid

1. health.cw.
2. (health or health behaviour or health beliefs or health care or health care costs or health care utilization or health care workers or health centres or health claims or health clinics or health clubs or health education or health foods or health hazards or health impact assessment or health indicators or health inequalities or health inspections or health insurance or health maintenance organizations or health policy or health programmes or health programs or health promotion or health protection or health resorts or health services or health tourism).de.
3. vv000.xc.
4. or/1-3
5. externalities/ or behavior modification/ or social interaction/ or peer influence/
6. ((herd effect*1 AND (vaccin* OR immunis* OR immuniz*)) or (cascade adj3 effect*) or (indirect adj3 "causal effect*") or (indirect adj3 protect*) or (infectious adj3 effect*) or spillover* or "spill over*" or externalit* or (interference and (vaccin* or immuniz* or immunis*)) or (interference and causal) or (peer* adj3 (effect* or influenc*)) or (("social network*" or "social* interact*") and (connect* or diffus* or spread* or (spill* adj over) or externalit* or contag* or infect* or transmis* or transmit*)) or "herd immunity" or "herd protect*" or "unexpected benefit*" or "unexpected effect*" or (("behavio* adj3 chang*" or "behavio* adj3 modif*") and (connect* or diffus* or spread* or (spill* adj over) or externalit* or contag* or infect* or transmis* or transmit*))).mp.
7. developing countries/ or least developed countries/
8. (Africa or Asia or Caribbean or West Indies or South America or Latin America or Central America).ti,ab,hw.
9. (Afghanistan or "Agalega Islands" or Albania or Algeria or Angola or Antigua or Argentina or Armenia or Armenian or Azerbaijan or Bangladesh or Barbuda or Basutoland or Belarus or Belize or Belorussia or Belorussian or Benin or Bhutan or Bolivia or Bosnia or Botswana or Brasil or Brazil or Bulgaria or "Burkina Faso" or "Burkina Fasso" or Burma or Burundi or Byelarus or Byelorussian or Cambodia or Cameron or Camerons or Cameroon or Cameroons or "Cape Verde" or "Central African Republic" or Ceylon or Chad or Chile or China or Colombia or Comores or "Comoro Islands" or Comoros or Congo or "Costa Rica" or "Cote d'Ivoire" or Cuba or Djibouti or Dominica or "Dominican Republic" or "East Timor" or "East Timur" or Ecuador or Egypt or "El Salvador" or Eritrea or Ethiopia or Fiji or "French Somaliland" or Gabon or "Gabonese Republic" or Gambia or Gaza or "Georgia Republic" or "Georgian Republic" or Ghana or "Gold Coast" or Grenada or Grenadines or Guatemala or Guiana or Guinea or "Guinea Bissau" or Guyana or Haiti or Hercegovina or Herzegovina or Honduras or Ifni or India or "Indian Ocean" or Indonesia or Iran or Iraq or "Ivory Coast" or Jamahiriya or Jamahiryria or Jamaica or Jordan or Kampuchea or Kazakh or Kazakhstan or Kenya or "Khmer Republic" or Kirghiz or Kirghizia or Kirgizstan or Kiribati or Korea or Kosovo or "Kyrgyz Republic" or Kyrgyzstan or "Lao PDR" or Laos or Latvia or Lebanon or Lesotho or Liberia or Libia or Libya or Lithuania or Macedonia or Madagascar or "Malagasy Republic" or

Malawi or Malay or Malaya or Malaysia or Maldives or Mali or "Marshall Islands" or Mauritania or Mauritius or Mayotte or Melanesia or Mexico or Micronesia or Mocambique or Moldova or Moldovia or Moldovian or Mongolia or Montenegro or Montenegro or Morocco or Mozambique or Myanma or Myanmar or Namibia or "Navigator Island" or "Navigator Islands" or Nepal or "New Caledonia" or "New Hebrides" or Nicaragua or Nicaragua or Niger or Nigeria or "Northern Mariana Islands" or Nyasaland or Pakistan or Palau or Palestine or Panama or "Papua New Guinea" or Paraguay or Peru or Philipines or Philippines or Phillipines or Phillippines or Principe or Rhodesia or Romania or Roumania or Ruanda or Rumania or Russia or Russian or "Russian Federation" or Rwanda or Sabah or "Saint Lucia" or "Saint Vincent" or Samoa or "Samoan Islands" or "Sao Tome" or Sarawak or Senegal or Serbia or Seychelles or "Sierra Leone" or "Solomon Islands" or Somalia or "South Africa" or "South Sudan" or "Soviet Union" or "Sri Lanka" or "St Lucia" or "St Vincent" or Sudan or Surinam or Suriname or Swaziland or Syria or Syrian or Tadjikistan or Tadzhik or Tadzhikistan or Tajikistan or Tanzania or Thailand or "Timor Leste" or Togo or "Togolese Republic" or Tonga or Tunisia or Turkey or Turkmen or Turkmenistan or Tuvalu or Uganda or Ukraine or "Union of Soviet Socialist Republics" or "United Arab Republic" or "Upper Volta" or Uruguay or Urundi or USSR or Uzbek or Uzbekistan or Vanuatu or Venezuela or "Viet Nam" or Vietnam or "West Bank" or "Western Sahara" or Yemen or Yugoslavia or Zaire or Zambia or Zimbabwe).tw,hw,sh.

10. ((developing or less* developed or under developed or underdeveloped or middle income or low* income or underserved or under served or deprived or poor*) adj (countr* or nation? or population? or world or state*)).ti,ab.
11. ((developing or less* developed or under developed or underdeveloped or middle income or low* income) adj (economy or economies)).ti,ab.
12. (low* adj (gdp or gnp or gross domestic or gross national)).tw.
13. (low adj3 middle adj3 countr*).tw.
14. (lmic or lmics or third world or lami countr*).tw.
15. transitional countr*.tw.
16. 7 or 8 or 9 or 10 or 11 or 12 or 13 or 14 or 15
17. 4 and (5 or 6) and 16
18. Limit 17 to (chinese or english or french or spanish)

Cochrane Library

- **First, for each item below, go to “Medical Terms (MeSH)” and click “Add to Search Manager” for the following terms with the explode option. This will create the first X entries in the list below.**
    - “Immunity, Herd”, Africa, South America, Mexico, Caribbean Region, Central America, Gulf of Mexico, Latin America, Asia, Pacific Islands, Developing Countries
- **Go to search manager to conduct the remaining searches.**

1. MeSH descriptor: [Immunity, Herd] explode all trees
2. MeSH descriptor: [Africa] explode all trees
3. MeSH descriptor: [South America] explode all trees
4. MeSH descriptor: [Mexico] explode all trees
5. MeSH descriptor: [Caribbean Region] explode all trees

6. MeSH descriptor: [Central America] explode all trees
7. MeSH descriptor: [Gulf of Mexico] explode all trees
8. MeSH descriptor: [Latin America] explode all trees
9. MeSH descriptor: [Asia] explode all trees
10. MeSH descriptor: [Pacific Islands] explode all trees
11. MeSH descriptor: [Developing Countries] explode all trees
12. ((herd effect*1 AND (vaccin* OR immunis* OR immuniz*)) OR (cascade adj3 effect*1) or (indirect adj3 causal effect*) or indirect* adj3 protect* or (infectiousness adj3 effect*) or spillover* or (spill*1 over) or externalit* or (interference and (vaccin* OR immunis* OR immuniz*)) OR (vaccine efficacy AND infectious*) OR (interference AND causal) or (peer* adj3 (effect* or influenc*)) or ((social network* or social* interact*) and (connect* or diffus* or spread* or (spill* adj over) or externalit* or contag* or infect* or transmis* or transmit*)) or herd immunity or herd protect* or unexpected benefit* or unexpected effect* or ((behavio* ad3 chang*or behavio* adj3 modif*) and (connect* or diffus* or spread* or (spill* adj over) or externalit* or contag* or infect* or transmis* or transmit*))).ti,ab
13. Developing Countries.kw
14. (Africa or Asia or Caribbean or West Indies or South America or Latin America or Central America):ti,ab,kw
15. (Afghanistan or "Agalega Islands" or Albania or Algeria or Angola or Antigua or Argentina or Armenia or Armenian or Azerbaijan or Bangladesh or Barbuda or Basutoland or Belarus or Belize or Belorussia or Belorussian or Benin or Bhutan or Bolivia or Bosnia or Botswana or Brasil or Brazil or Bulgaria or "Burkina Faso" or "Burkina Fasso" or Burma or Burundi or Byelarus or Byelorussian or Cambodia or Cameron or Camerons or Cameroon or Cameroons or "Cape Verde" or "Central African Republic" or Ceylon or Chad or Chile or China or Colombia or Comores or "Comoro Islands" or Comoros or Congo or "Costa Rica" or "Cote d'Ivoire" or Cuba or Djibouti or Dominica or "Dominican Republic" or "East Timor" or "East Timur" or Ecuador or Egypt or "El Salvador" or Eritrea or Ethiopia or Fiji or "French Somaliland" or Gabon or "Gabonese Republic" or Gambia or Gaza or "Georgia Republic" or "Georgian Republic" or Ghana or "Gold Coast" or Grenada or Grenadines or Guatemala or Guiana or Guinea or "Guinea Bissau" or Guyana or Haiti or Hercegovina or Herzegovina or Honduras or Ifni or India or "Indian Ocean" or Indonesia or Iran or Iraq or "Ivory Coast" or Jamahiriya or Jamahiryria or Jamaica or Jordan or Kampuchea or Kazakh or Kazakhstan or Kenya or "Khmer Republic" or Kirghiz or Kirghizia or Kirgizstan or Kiribati or Korea or Kosovo or "Kyrgyz Republic" or Kyrgyzstan or "Lao PDR" or Laos or Latvia or Lebanon or Lesotho or Liberia or Libia or Libya or Lithuania or Macedonia or Madagascar or "Malagasy Republic" or Malawi or Malay or Malaya or Malaysia or Maldives or Mali or "Marshall Islands" or Mauritania or Mauritius or Mayotte or Melanesia or Mexico or Micronesia or Mocambique or Moldova or Moldovia or Moldovian or Mongolia or Montenegro or Montenegro or Morocco or Mozambique or Myanma or Myanmar or Namibia or "Navigator Island" or "Navigator Islands" or Nepal or "New Caledonia" or "New Hebrides" or Nicaragua or Nicaragua or Niger or Nigeria or "Northern Mariana Islands" or Nyasaland or Pakistan or Palau or Palestine or Panama or "Papua New Guinea" or Paraguay or Peru or Philipines or Philippines or Phillipines or Phillippines or Principe or Rhodesia or Romania or Roumania or Ruanda or Rumania or Russia

or Russian or "Russian Federation" or Rwanda or Sabah or "Saint Lucia" or "Saint Vincent" or Samoa or "Samoan Islands" or "Sao Tome" or Sarawak or Senegal or Serbia or Seychelles or "Sierra Leone" or "Solomon Islands" or Somalia or "South Africa" or "South Sudan" or "Soviet Union" or "Sri Lanka" or "St Lucia" or "St Vincent" or Sudan or Surinam or Suriname or Swaziland or Syria or Syrian or Tadjikistan or Tadzhik or Tadzhikistan or Tajikistan or Tanzania or Thailand or "Timor Leste" or Togo or "Togolese Republic" or Tonga or Tunisia or Turkey or Turkmen or Turkmenistan or Tuvalu or Uganda or Ukraine or "Union of Soviet Socialist Republics" or "United Arab Republic" or "Upper Volta" or Uruguay or Urundi or USSR or Uzbek or Uzbekistan or Vanuatu or Venezuela or "Viet Nam" or Vietnam or "West Bank" or "Western Sahara" or Yemen or Yugoslavia or Zaire or Zambia or Zimbabwe):kw,ti,ab

16. ((developing or less* developed or "under developed" or underdeveloped or "middle income" or low* income or underserved or "under served" or deprived or poor*) near (countr* or nation* or population* or world or state*)):ti,ab
17. ((developing or less* developed or “under developed” or underdeveloped or “middle income” or low* income) near (economy or economies)):ti,ab
18. (low* near (gdp or gnp or gross domestic or gross national)):ti,ab
19. (low near/3 middle near/3 countr*):ti,ab
20. (lmic or lmics or "third world" or "lami countr*"):ti,ab
21. “transitional countr*”:ti,ab
22. (1 or 12) and ((or/2-11) or (or/13-21))
23. Limit 22 to (chinese or english or french or spanish)

Econlit

1. ((((herd effect*1 and (vaccin* or immunis* or immuniz*)) or (cascade adj3 effect*1) or (indirect adj3 causal effect*) or indirect*) adj3 protect*) or (infectiousness adj3 effect*) or spillover* or spill*1 over or externalit* or (interference and (vaccin* or immunis* or immuniz*)) or (vaccine efficacy and infectious*) or (interference and causal) or (peer* adj3 (effect* or influenc*)) or ((social network* or social* interact*) and (connect* or diffus* or spread* or (spill* adj over) or externalit* or contag* or infect* or transmis* or transmit*)) or herd immunity or herd protect* or unexpected benefit* or unexpected effect* or ((behavio* adj3 chang*or behavio* adj3 modif*) and (connect* or diffus* or spread* or (spill* adj over) or externalit* or contag* or infect* or transmis* or transmit*))).ti,ab.
2. Externalities.hw.
3. 1 or 2
4. (Africa or Asia or Caribbean or West Indies or South America or Latin America or Central America).hw,ti,ab.
5. (Afghanistan or "Agalega Islands" or Albania or Algeria or Angola or Antigua or Argentina or Armenia or Armenian or Azerbaijan or Bangladesh or Barbuda or Basutoland or Belarus or Belize or Belorussia or Belorussian or Benin or Bhutan or Bolivia or Bosnia or Botswana or Brasil or Brazil or Bulgaria or "Burkina Faso" or "Burkina Fasso" or Burma or Burundi or Byelarus or Byelorussian or Cambodia or Cameron or Camerons or Cameroon or Cameroons or "Cape Verde" or "Central African Republic" or Ceylon or Chad or Chile or China or Colombia or Comores or "Comoro Islands" or Comoros or Congo or "Costa Rica" or "Cote d'Ivoire" or Cuba or

Djibouti or Dominica or "Dominican Republic" or "East Timor" or "East Timur" or Ecuador or Egypt or "El Salvador" or Eritrea or Ethiopia or Fiji or "French Somaliland" or Gabon or "Gabonese Republic" or Gambia or Gaza or "Georgia Republic" or "Georgian Republic" or Ghana or "Gold Coast" or Grenada or Grenadines or Guatemala or Guiana or Guinea or "Guinea Bissau" or Guyana or Haiti or Hercegovina or Herzegovina or Honduras or Ifni or India or "Indian Ocean" or Indonesia or Iran or Iraq or "Ivory Coast" or Jamahiriya or Jamahiryria or Jamaica or Jordan or Kampuchea or Kazakh or Kazakhstan or Kenya or "Khmer Republic" or Kirghiz or Kirghizia or Kirgizstan or Kiribati or Korea or Kosovo or "Kyrgyz Republic" or Kyrgyzstan or "Lao PDR" or Laos or Latvia or Lebanon or Lesotho or Liberia or Libia or Libya or Lithuania or Macedonia or Madagascar or "Malagasy Republic" or Malawi or Malay or Malaya or Malaysia or Maldives or Mali or "Marshall Islands" or Mauritania or Mauritius or Mayotte or Melanesia or Mexico or Micronesia or Mocambique or Moldova or Moldovia or Moldovian or Mongolia or Montenegro or Montenegro or Morocco or Mozambique or Myanma or Myanmar or Namibia or "Navigator Island" or "Navigator Islands" or Nepal or "New Caledonia" or "New Hebrides" or Nicaragua or Nicaragua or Niger or Nigeria or "Northern Mariana Islands" or Nyasaland or Pakistan or Palau or Palestine or Panama or "Papua New Guinea" or Paraguay or Peru or Philipines or Philippines or Phillipines or Phillippines or Principe or Rhodesia or Romania or Roumania or Ruanda or Rumania or Russia or Russian or "Russian Federation" or Rwanda or Sabah or "Saint Lucia" or "Saint Vincent" or Samoa or "Samoan Islands" or "Sao Tome" or Sarawak or Senegal or Serbia or Seychelles or "Sierra Leone" or "Solomon Islands" or Somalia or "South Africa" or "South Sudan" or "Soviet Union" or "Sri Lanka" or "St Lucia" or "St Vincent" or Sudan or Surinam or Suriname or Swaziland or Syria or Syrian or Tadjikistan or Tadzhik or Tadzhikistan or Tajikistan or Tanzania or Thailand or "Timor Leste" or Togo or "Togolese Republic" or Tonga or Tunisia or Turkey or Turkmen or Turkmenistan or Tuvalu or Uganda or Ukraine or "Union of Soviet Socialist Republics" or "United Arab Republic" or "Upper Volta" or Uruguay or Urundi or USSR or Uzbek or Uzbekistan or Vanuatu or Venezuela or "Viet Nam" or Vietnam or "West Bank" or "Western Sahara" or Yemen or Yugoslavia or Zaire or Zambia or Zimbabwe).hw,ti,ab.

6. (developing country or developing countries).hw.
7. ((developing or less* developed or under developed or underdeveloped or middle income or low* income or underserved or under served or deprived or poor*) adj (countr* or nation? or population? or world or state*)).ti,ab.
8. ((developing or less* developed or under developed or underdeveloped or middle income or low* income) adj (economy or economies)).ti,ab.
9. (low* adj (gdp or gnp or gross domestic or gross national)).ti,ab.
10. (low adj3 middle adj3 countr*).ti,ab.
11. (lmic or lmics or third world or lami countr*).ti,ab.
12. transitional countr*.ti,ab.
13. or/4-12
14. 3 and 13
15. (chinese or english or french or spanish).lg.
16. 14 and 15

Embase

1. herd immunity/ or behavior change/ or *"work home spillover"/ or social network/ or peer group/
2. ((((herd effect*1 and (vaccin* or immunis* or immuniz*)) or (cascade adj3 effect*1) or (indirect adj3 causal effect*) or indirect*) adj3 protect*) or (infectiousness adj3 effect*) or spillover* or spill*1 over or externalit* or (interference and (vaccin* or immunis* or immuniz*)) or (vaccine efficacy and infectious*) or (interference and causal) or (peer* adj3 (effect* or influenc*)) or ((social network* or social* interact*) and (connect* or diffus* or spread* or (spill* adj over) or externalit* or contag* or infect* or transmis* or transmit*)) or herd immunity or herd protect* or unexpected benefit* or unexpected effect* or ((behavio* adj3 chang*or behavio* adj3 modif*) and (connect* or diffus* or spread* or (spill* adj over) or externalit* or contag* or infect* or transmis* or transmit*))).ti,ab.
3. or/1-2
4. Developing Country.sh.
5. (Africa or Asia or Caribbean or West Indies or South America or Latin America or Central America).hw,ti,ab,cp.
6. (Afghanistan or "Agalega Islands" or Albania or Algeria or Angola or Antigua or Argentina or Armenia or Armenian or Azerbaijan or Bangladesh or Barbuda or Basutoland or Belarus or Belize or Belorussia or Belorussian or Benin or Bhutan or Bolivia or Bosnia or Botswana or Brasil or Brazil or Bulgaria or "Burkina Faso" or "Burkina Fasso" or Burma or Burundi or Byelarus or Byelorussian or Cambodia or Cameron or Camerons or Cameroon or Cameroons or "Cape Verde" or "Central African Republic" or Ceylon or Chad or Chile or China or Colombia or Comores or "Comoro Islands" or Comoros or Congo or "Costa Rica" or "Cote d'Ivoire" or Cuba or Djibouti or Dominica or "Dominican Republic" or "East Timor" or "East Timur" or Ecuador or Egypt or "El Salvador" or Eritrea or Ethiopia or Fiji or "French Somaliland" or Gabon or "Gabonese Republic" or Gambia or Gaza or "Georgia Republic" or "Georgian Republic" or Ghana or "Gold Coast" or Grenada or Grenadines or Guatemala or Guiana or Guinea or "Guinea Bissau" or Guyana or Haiti or Hercegovina or Herzegovina or Honduras or Ifni or India or "Indian Ocean" or Indonesia or Iran or Iraq or "Ivory Coast" or Jamahiriya or Jamahiryria or Jamaica or Jordan or Kampuchea or Kazakh or Kazakhstan or Kenya or "Khmer Republic" or Kirghiz or Kirghizia or Kirgizstan or Kiribati or Korea or Kosovo or "Kyrgyz Republic" or Kyrgyzstan or "Lao PDR" or Laos or Latvia or Lebanon or Lesotho or Liberia or Libia or Libya or Lithuania or Macedonia or Madagascar or "Malagasy Republic" or Malawi or Malay or Malaya or Malaysia or Maldives or Mali or "Marshall Islands" or Mauritania or Mauritius or Mayotte or Melanesia or Mexico or Micronesia or Mocambique or Moldova or Moldovia or Moldovian or Mongolia or Montenegro or Montenegro or Morocco or Mozambique or Myanma or Myanmar or Namibia or "Navigator Island" or "Navigator Islands" or Nepal or "New Caledonia" or "New Hebrides" or Nicaragua or Nicaragua or Niger or Nigeria or "Northern Mariana Islands" or Nyasaland or Pakistan or Palau or Palestine or Panama or "Papua New Guinea" or Paraguay or Peru or Philipines or Philippines or Phillipines or Phillippines or Principe or Rhodesia or Romania or Roumania or Ruanda or Rumania or Russia or Russian or "Russian Federation" or Rwanda or Sabah or "Saint Lucia" or "Saint Vincent" or Samoa or "Samoan Islands" or "Sao Tome" or Sarawak or Senegal or

Serbia or Seychelles or "Sierra Leone" or "Solomon Islands" or Somalia or "South Africa" or "South Sudan" or "Soviet Union" or "Sri Lanka" or "St Lucia" or "St Vincent" or Sudan or Surinam or Suriname or Swaziland or Syria or Syrian or Tadjikistan or Tadzhik or Tadzhikistan or Tajikistan or Tanzania or Thailand or "Timor Leste" or Togo or "Togolese Republic" or Tonga or Tunisia or Turkey or Turkmen or Turkmenistan or Tuvalu or Uganda or Ukraine or "Union of Soviet Socialist Republics" or "United Arab Republic" or "Upper Volta" or Uruguay or Urundi or USSR or Uzbek or Uzbekistan or Vanuatu or Venezuela or "Viet Nam" or Vietnam or "West Bank" or "Western Sahara" or Yemen or Yugoslavia or Zaire or Zambia or Zimbabwe).hw,ti,ab,cp.

7. ((developing or less* developed or under developed or underdeveloped or middle income or low* income or underserved or under served or deprived or poor*) adj (countr* or nation? or population? or world or state*)).ti,ab.
8. ((developing or less* developed or under developed or underdeveloped or middle income or low* income) adj (economy or economies)).ti,ab.
9. (low* adj (gdp or gnp or gross domestic or gross national)).ti,ab.
10. (low adj3 middle adj3 countr*).ti,ab.
11. (lmic or lmics or third world or lami countr*).ti,ab.
12. transitional countr*.ti,ab.
13. or/4-12
14. 3 and 13
15. limit 14 to (chinese or english or french or spanish)
16. Animal/ not (Animal/ and Human/)
17. 15 not 16

ERIC/ProQuest

1. su("Peer influence" OR "community influence" OR “environmental influences” OR “social influences” OR “social distance” OR “social experience” OR “behavior change” OR “behavior modification” OR “economic development”)
2. ti,ab((“herd effect*1” AND (vaccin* OR immunis* OR immuniz*)) OR (cascade near/3 effect*1) or (indirect near/3 “causal effect*”) or (indirect* near/3 protect*) or (infectiousness near/3 effect*) or spillover* or (spill*1 over) or externalit* or (interference and (vaccin* OR immunis* OR immuniz*)) OR (“vaccine efficacy” AND infectious*) OR (interference AND causal) or (peer* near/3 (effect* or influenc*)) or ((“social network*” or “social* interact*”) and (connect* or diffus* or spread* or (spill* adj over) or externalit* or contag* or infect* or transmis* or transmit*)) or “herd immunity” or “herd protect*” or “unexpected benefit*” or “unexpected effect*” or ((behavio* ad3 chang*or behavio* adj3 modif*) and (connect* or diffus* or spread* or (spill* adj over) or externalit* or contag* or infect* or transmis* or transmit*)))
3. SU.EXACT.EXPLODE("Health" OR "child health" OR "health programs" OR "health promotion" OR "community health services" OR "health behavior" OR "Comprehensive School Health Education" OR "health activities")
4. la(english or spanish or chinese or french)
5. su(“developing nations” OR “developing nations AND economic development”)
6. ti,ab,su,loc(Africa or Asia or Caribbean or “West Indies” or “South America” or “Latin America” or “Central America”)((ti,ab((Afghanistan or "Agalega Islands" or Albania or Algeria or Angola or Antigua or Argentina or Armenia or Armenian or Azerbaijan or

Bangladesh or Barbuda or Basutoland or Belarus or Belize or Belorussia or Belorussian or Benin or Bhutan or Bolivia or Bosnia or Botswana or Brasil or Brazil or Bulgaria or "Burkina Faso" or "Burkina Fasso" or Burma or Burundi or Byelarus or Byelorussian or Cambodia or Cameron or Camerons or Cameroon or Cameroons or "Cape Verde" or "Central African Republic" or Ceylon or Chad or Chile or China or Colombia or Comores or "Comoro Islands" or Comoros or Congo or "Costa Rica" or "Cote d'Ivoire" or Cuba or Djibouti or Dominica or "Dominican Republic" or "East Timor" or "East Timur" or Ecuador or Egypt or "El Salvador" or Eritrea or Ethiopia or Fiji or "French Somaliland" or Gabon or "Gabonese Republic" or Gambia or Gaza or "Georgia Republic" or "Georgian Republic" or Ghana or "Gold Coast" or Grenada or Grenadines or Guatemala or Guiana or Guinea or "Guinea Bissau" or Guyana or Haiti or Hercegovina or Herzegovina or Honduras or Ifni or India or "Indian Ocean" or Indonesia or Iran or Iraq or "Ivory Coast" or Jamahiriya or Jamahiryria or Jamaica or Jordan or Kampuchea or Kazakh or Kazakhstan or Kenya or "Khmer Republic" or Kirghiz or Kirghizia or Kirgizstan or Kiribati or Korea or Kosovo or "Kyrgyz Republic" or Kyrgyzstan or "Lao PDR" or Laos or Latvia or Lebanon or Lesotho or Liberia or Libia or Libya or Lithuania or Macedonia or Madagascar or "Malagasy Republic" or Malawi or Malay or Malaya or Malaysia or Maldives or Mali or "Marshall Islands" or Mauritania or Mauritius or Mayotte or Melanesia or Mexico or Micronesia or Mocambique or Moldova or Moldovia or Moldovian or Mongolia or Montenegro or Montenegro or Morocco or Mozambique or Myanma or Myanmar or Namibia or "Navigator Island" or "Navigator Islands" or Nepal or "New Caledonia" or "New Hebrides" or Nicaragua or Nicaragua or Niger or Nigeria or "Northern Mariana Islands" or Nyasaland or Pakistan or Palau or Palestine or Panama or "Papua New Guinea" or Paraguay or Peru or Philipines or Philippines or Phillipines or Phillippines or Principe or Rhodesia or Romania or Roumania or Ruanda or Rumania or Russia or Russian or "Russian Federation" or Rwanda or Sabah or "Saint Lucia" or "Saint Vincent" or Samoa or "Samoan Islands" or "Sao Tome" or Sarawak or Senegal or Serbia or Seychelles or "Sierra Leone" or "Solomon Islands" or Somalia or "South Africa" or "South Sudan" or "Soviet Union" or "Sri Lanka" or "St Lucia" or "St Vincent" or Sudan or Surinam or Suriname or Swaziland or Syria or Syrian or Tadjikistan or Tadzhik or Tadzhikistan or Tajikistan or Tanzania or Thailand or "Timor Leste" or Togo or "Togolese Republic" or Tonga or Tunisia or Turkey or Turkmen or Turkmenistan or Tuvalu or Uganda or Ukraine or "Union of Soviet Socialist Republics" or "United Arab Republic" or "Upper Volta" or Uruguay or Urundi or USSR or Uzbek or Uzbekistan or Vanuatu or Venezuela or "Viet Nam" or Vietnam or "West Bank" or "Western Sahara" or Yemen or Yugoslavia or Zaire or Zambia or Zimbabwe))))

7. ti,ab((developing or "less* developed" or "under developed" or underdeveloped or "middle income" or "low* income" or underserved or "under served" or deprived or poor*) near (countr* or nation*1 or population*1 or world or state*))
8. ti,ab((developing or “less* developed” or “under developed” or underdeveloped or “middle income” or “low* income”) near (economy or economies))
9. ti,ab(low* near (gdp or gnp or “gross domestic” or “gross national”))
10. ti,ab(low near/3 middle near/3 countr*)
11. ti,ab(LMIC OR LMICS OR "third world" OR "LAMI countr*")
12. ti,ab(transitional countr*)
13. (1 or 2) and 3 and 4 and (5 or 6 or 7 or 8 or 9 or 10 or 11 or 12 or 13 or 14)

Global Health/Ovid

1. ((herd effect*1 AND (vaccin* OR immunis* OR immuniz*)) or (cascade adj3 effect*) or (indirect adj3 "causal effect*") or (indirect adj3 protect*) or (infectious adj3 effect*) or spillover* or "spill over*" or externalit* or (interference and (vaccin* or immuniz* or immunis*)) or (interference and causal) or (peer* adj3 (effect* or influenc*)) or (("social network*" or "social* interact*") and (connect* or diffus* or spread* or (spill* adj over) or externalit* or contag* or infect* or transmis* or transmit*)) or "herd immunity" or "herd protect*" or "unexpected benefit*" or "unexpected effect*" or (("behavio* adj3 chang*" or "behavio* adj3 modif*") and (connect* or diffus* or spread* or (spill* adj over) or externalit* or contag* or infect* or transmis* or transmit*))).mp. [mp=abstract, title, original title, broad terms, heading words]
2. externalities/ or behavior modification/ or social interaction/ or peer influence/
3. developing countries/ or least developed countries/
4. (Africa or Asia or Caribbean or West Indies or South America or Latin America or Central America).ti,ab,hw.
5. (Afghanistan or "Agalega Islands" or Albania or Algeria or Angola or Antigua or Argentina or Armenia or Armenian or Azerbaijan or Bangladesh or Barbuda or Basutoland or Belarus or Belize or Belorussia or Belorussian or Benin or Bhutan or Bolivia or Bosnia or Botswana or Brasil or Brazil or Bulgaria or "Burkina Faso" or "Burkina Fasso" or Burma or Burundi or Byelarus or Byelorussian or Cambodia or Cameron or Camerons or Cameroon or Cameroons or "Cape Verde" or "Central African Republic" or Ceylon or Chad or Chile or China or Colombia or Comores or "Comoro Islands" or Comoros or Congo or "Costa Rica" or "Cote d'Ivoire" or Cuba or Djibouti or Dominica or "Dominican Republic" or "East Timor" or "East Timur" or Ecuador or Egypt or "El Salvador" or Eritrea or Ethiopia or Fiji or "French Somaliland" or Gabon or "Gabonese Republic" or Gambia or Gaza or "Georgia Republic" or "Georgian Republic" or Ghana or "Gold Coast" or Grenada or Grenadines or Guatemala or Guiana or Guinea or "Guinea Bissau" or Guyana or Haiti or Hercegovina or Herzegovina or Honduras or Ifni or India or "Indian Ocean" or Indonesia or Iran or Iraq or "Ivory Coast" or Jamahiriya or Jamahiryria or Jamaica or Jordan or Kampuchea or Kazakh or Kazakhstan or Kenya or "Khmer Republic" or Kirghiz or Kirghizia or Kirgizstan or Kiribati or Korea or Kosovo or "Kyrgyz Republic" or Kyrgyzstan or "Lao PDR" or Laos or Latvia or Lebanon or Lesotho or Liberia or Libia or Libya or Lithuania or Macedonia or Madagascar or "Malagasy Republic" or Malawi or Malay or Malaya or Malaysia or Maldives or Mali or "Marshall Islands" or Mauritania or Mauritius or Mayotte or Melanesia or Mexico or Micronesia or Mocambique or Moldova or Moldovia or Moldovian or Mongolia or Montenegro or Montenegro or Morocco or Mozambique or Myanma or Myanmar or Namibia or "Navigator Island" or "Navigator Islands" or Nepal or "New Caledonia" or "New Hebrides" or Nicaragua or Nicaragua or Niger or Nigeria or "Northern Mariana Islands" or Nyasaland or Pakistan or Palau or Palestine or Panama or "Papua New Guinea" or Paraguay or Peru or Philipines or Philippines or Phillipines or Phillippines or Principe or Rhodesia or Romania or Roumania or Ruanda or Rumania or Russia or Russian or "Russian Federation" or Rwanda or Sabah or "Saint Lucia" or "Saint Vincent" or Samoa or "Samoan Islands" or "Sao Tome" or Sarawak or Senegal or Serbia or Seychelles or "Sierra Leone" or "Solomon Islands" or Somalia or "South Africa" or "South Sudan" or "Soviet Union" or "Sri Lanka" or "St Lucia" or "St Vincent"

or Sudan or Surinam or Suriname or Swaziland or Syria or Syrian or Tadjikistan or Tadzhik or Tadzhikistan or Tajikistan or Tanzania or Thailand or "Timor Leste" or Togo or "Togolese Republic" or Tonga or Tunisia or Turkey or Turkmen or Turkmenistan or Tuvalu or Uganda or Ukraine or "Union of Soviet Socialist Republics" or "United Arab Republic" or "Upper Volta" or Uruguay or Urundi or USSR or Uzbek or Uzbekistan or Vanuatu or Venezuela or "Viet Nam" or Vietnam or "West Bank" or "Western Sahara" or Yemen or Yugoslavia or Zaire or Zambia or Zimbabwe).tw,hw,sh.

6. ((developing or less* developed or under developed or underdeveloped or middle income or low* income or underserved or under served or deprived or poor*) adj (countr* or nation? or population? or world or state*)).ti,ab.
7. ((developing or less* developed or under developed or underdeveloped or middle income or low* income) adj (economy or economies)).ti,ab.
8. (low* adj (gdp or gnp or gross domestic or gross national)).tw.
9. (low adj3 middle adj3 countr*).tw.
10. (lmic or lmics or third world or lami countr*).tw.
11. transitional countr*.tw.
12. (1 or 2) and or/3-11
13. Limit 12 to (chinese or english or french or spanish)

IBSS/Proquest

1. su("externalities" OR "group influence" OR "social influence" OR “interpersonal influence” OR “social contagion“ OR "social distance" OR "social forces" OR "social interaction" OR "social network" OR "collective behavior" OR "collective behaviour" OR "economic behaviour" OR "group behaviour" OR "social behaviour" OR “economic development” OR “economic development projects” OR ”economic & social development” OR ”economic and social development” OR ”socio-economic development”)
2. ti,ab((“herd effect*1” AND (vaccin* OR immunis* OR immuniz*)) OR (cascade near/3 effect*1) or (indirect near/3 “causal effect*”) or (indirect* near/3 protect*) or (infectiousness near/3 effect*) or spillover* or (spill*1 over) or externalit* or (interference and (vaccin* OR immunis* OR immuniz*)) OR (“vaccine efficacy” AND infectious*) OR (interference AND causal) or (peer* near/3 (effect* or influenc*)) or ((“social network*” or “social* interact*”) and (connect* or diffus* or spread* or (spill* adj over) or externalit* or contag* or infect* or transmis* or transmit*)) or “herd immunity” or “herd protect*” or “unexpected benefit*” or “unexpected effect*” or ((behavio* ad3 chang*or behavio* adj3 modif*) and (connect* or diffus* or spread* or (spill* adj over) or externalit* or contag* or infect* or transmis* or transmit*)))
3. SU.EXPLODE("Addiction" OR "AIDS" OR "Alcoholism" OR "Alzheimer's disease" OR "Anorexia nervosa" OR "Cancer" OR "Child health" OR "Congenital deformities" OR "Cranial deformations" OR "Cytology" OR "Deformations" OR "Depression" OR "Dietary disorders" OR "Diseases" OR "Drug addiction" OR "Eating disorders" OR "Epidemics" OR "Epilepsy" OR "Etiology" OR "Food safety" OR "Health" OR "Health aid" OR "Health care" OR "Health centres" OR "Health economics" OR "Health education" OR "Health expenditure" OR "Health inequality" OR "Heart disease" OR "Health services" OR "Hepatitis" OR "HIV" OR "Hospices" OR "Hospital services" OR "Hygiene" OR "Illness" OR "Immunization" OR "Injuries" OR "Leprosy" OR

"Madness" OR "Malaria" OR "Measles" OR "Medical care" OR "Medical personnel" OR "Medical sociology" OR "Medical treatment" OR "Mental health" OR "Mental hygiene" OR "Mental illness" OR "Midwives" OR "Neuroses" OR "Occupational diseases" OR "Occupational health" OR "Paramedical personnel" OR "Patients" OR "Personality disorders" OR "Physiotherapy" OR "Primary health care" OR "Private health care" OR "Psychoses" OR "Public health" OR "Reproductive health" OR "Schizophrenia" OR "Sexual health" OR "Sexually transmitted diseases" OR "Smallpox" OR "Social hygiene" OR "Transplants" OR "Trauma" OR "Trypanosomiasis" OR "Tuberculosis" OR "Women's health" OR "Yellow fever")

4. la(english or spanish or chinese or french)
5. su(“developing countries”)
6. ti,ab,su,loc(Africa or Asia or Caribbean or “West Indies” or “South America” or “Latin America” or “Central America”)
7. ((ti,ab((Afghanistan or "Agalega Islands" or Albania or Algeria or Angola or Antigua or Argentina or Armenia or Armenian or Azerbaijan or Bangladesh or Barbuda or Basutoland or Belarus or Belize or Belorussia or Belorussian or Benin or Bhutan or Bolivia or Bosnia or Botswana or Brasil or Brazil or Bulgaria or "Burkina Faso" or "Burkina Fasso" or Burma or Burundi or Byelarus or Byelorussian or Cambodia or Cameron or Camerons or Cameroon or Cameroons or "Cape Verde" or "Central African Republic" or Ceylon or Chad or Chile or China or Colombia or Comores or "Comoro Islands" or Comoros or Congo or "Costa Rica" or "Cote d'Ivoire" or Cuba or Djibouti or Dominica or "Dominican Republic" or "East Timor" or "East Timur" or Ecuador or Egypt or "El Salvador" or Eritrea or Ethiopia or Fiji or "French Somaliland" or Gabon or "Gabonese Republic" or Gambia or Gaza or "Georgia Republic" or "Georgian Republic" or Ghana or "Gold Coast" or Grenada or Grenadines or Guatemala or Guiana or Guinea or "Guinea Bissau" or Guyana or Haiti or Hercegovina or Herzegovina or Honduras or Ifni or India or "Indian Ocean" or Indonesia or Iran or Iraq or "Ivory Coast" or Jamahiriya or Jamahiryria or Jamaica or Jordan or Kampuchea or Kazakh or Kazakhstan or Kenya or "Khmer Republic" or Kirghiz or Kirghizia or Kirgizstan or Kiribati or Korea or Kosovo or "Kyrgyz Republic" or Kyrgyzstan or "Lao PDR" or Laos or Latvia or Lebanon or Lesotho or Liberia or Libia or Libya or Lithuania or Macedonia or Madagascar or "Malagasy Republic" or Malawi or Malay or Malaya or Malaysia or Maldives or Mali or "Marshall Islands" or Mauritania or Mauritius or Mayotte or Melanesia or Mexico or Micronesia or Mocambique or Moldova or Moldovia or Moldovian or Mongolia or Montenegro or Montenegro or Morocco or Mozambique or Myanma or Myanmar or Namibia or "Navigator Island" or "Navigator Islands" or Nepal or "New Caledonia" or "New Hebrides" or Nicaragua or Nicaragua or Niger or Nigeria or "Northern Mariana Islands" or Nyasaland or Pakistan or Palau or Palestine or Panama or "Papua New Guinea" or Paraguay or Peru or Philipines or Philippines or Phillipines or Phillippines or Principe or Rhodesia or Romania or Roumania or Ruanda or Rumania or Russia or Russian or "Russian Federation" or Rwanda or Sabah or "Saint Lucia" or "Saint Vincent" or Samoa or "Samoan Islands" or "Sao Tome" or Sarawak or Senegal or Serbia or Seychelles or "Sierra Leone" or "Solomon Islands" or Somalia or "South Africa" or "South Sudan" or "Soviet Union" or "Sri Lanka" or "St Lucia" or "St Vincent" or Sudan or Surinam or Suriname or Swaziland or Syria or Syrian or Tadjikistan or Tadzhik or Tadzhikistan or Tajikistan or Tanzania or Thailand or "Timor Leste" or Togo or "Togolese Republic" or Tonga or Tunisia or Turkey or Turkmen or

Turkmenistan or Tuvalu or Uganda or Ukraine or "Union of Soviet Socialist Republics" or "United Arab Republic" or "Upper Volta" or Uruguay or Urundi or USSR or Uzbek or Uzbekistan or Vanuatu or Venezuela or "Viet Nam" or Vietnam or "West Bank" or "Western Sahara" or Yemen or Yugoslavia or Zaire or Zambia or Zimbabwe))))

8. ti,ab((developing or "less* developed" or "under developed" or underdeveloped or "middle income" or "low* income" or underserved or "under served" or deprived or poor*) near (countr* or nation*1 or population*1 or world or state*))
9. ti,ab((developing or “less* developed” or “under developed” or underdeveloped or “middle income” or “low* income”) near (economy or economies))
10. ti,ab(low* near (gdp or gnp or “gross domestic” or “gross national”))
11. ti,ab(low near/3 middle near/3 countr*)
12. ti,ab(LMIC OR LMICS OR "third world" OR "LAMI countr*")
13. ti,ab("transitional countr*")
14. (1 or 2) and 3 and 4 and (5 or 6 or 7 or 8 or 9 or 10 or 11 or 12 or 13 or 14)

Medline/Ovid

1. Immunity, Herd/
2. ((herd effect*1 AND (vaccin* OR immunis* OR immuniz*)) OR (cascade adj3 effect*1) or (indirect adj3 causal effect*) or indirect* adj3 protect* or (infectiousness adj3 effect*) or spillover* or (spill*1 over) or externalit* or (interference and (vaccin* OR immunis* OR immuniz*)) OR (vaccine efficacy AND infectious*) OR (interference AND causal) or (peer* adj3 (effect* or influenc*)) or ((social network* or social* interact*) and (connect* or diffus* or spread* or (spill* adj over) or externalit* or contag* or infect* or transmis* or transmit*)) or herd immunity or herd protect* or unexpected benefit* or unexpected effect* or ((behavio* adj3 chang*or behavio* adj3 modif*) and (connect* or diffus* or spread* or (spill* adj over) or externalit* or contag* or infect* or transmis* or transmit*))).ti,ab
3. exp Africa/ OR exp South America/ OR exp Mexico/ OR exp Caribbean Region/ OR exp Central America/ OR exp Gulf of Mexico/ OR exp Latin America/ OR exp Asia/ OR exp Pacific Islands/
4. Developing countries/
5. Developing Countries.sh,kf.
6. (Africa or Asia or Caribbean or West Indies or South America or Latin America or Central America).hw,kf,ti,ab,cp.
7. (Afghanistan or "Agalega Islands" or Albania or Algeria or Angola or Antigua or Argentina or Armenia or Armenian or Azerbaijan or Bangladesh or Barbuda or Basutoland or Belarus or Belize or Belorussia or Belorussian or Benin or Bhutan or Bolivia or Bosnia or Botswana or Brasil or Brazil or Bulgaria or "Burkina Faso" or "Burkina Fasso" or Burma or Burundi or Byelarus or Byelorussian or Cambodia or Cameron or Camerons or Cameroon or Cameroons or "Cape Verde" or "Central African Republic" or Ceylon or Chad or Chile or China or Colombia or Comores or "Comoro Islands" or Comoros or Congo or "Costa Rica" or "Cote d'Ivoire" or Cuba or Djibouti or Dominica or "Dominican Republic" or "East Timor" or "East Timur" or Ecuador or Egypt or "El Salvador" or Eritrea or Ethiopia or Fiji or "French Somaliland" or Gabon or "Gabonese Republic" or Gambia or Gaza or "Georgia Republic" or "Georgian Republic" or Ghana or "Gold Coast" or Grenada or Grenadines or

Guatemala or Guiana or Guinea or "Guinea Bissau" or Guyana or Haiti or Hercegovina or Herzegovina or Honduras or Ifni or India or "Indian Ocean" or Indonesia or Iran or Iraq or "Ivory Coast" or Jamahiriya or Jamahiryria or Jamaica or Jordan or Kampuchea or Kazakh or Kazakhstan or Kenya or "Khmer Republic" or Kirghiz or Kirghizia or Kirgizstan or Kiribati or Korea or Kosovo or "Kyrgyz Republic" or Kyrgyzstan or "Lao PDR" or Laos or Latvia or Lebanon or Lesotho or Liberia or Libia or Libya or Lithuania or Macedonia or Madagascar or "Malagasy Republic" or Malawi or Malay or Malaya or Malaysia or Maldives or Mali or "Marshall Islands" or Mauritania or Mauritius or Mayotte or Melanesia or Mexico or Micronesia or Mocambique or Moldova or Moldovia or Moldovian or Mongolia or Montenegro or Montenegro or Morocco or Mozambique or Myanma or Myanmar or Namibia or "Navigator Island" or "Navigator Islands" or Nepal or "New Caledonia" or "New Hebrides" or Nicaragua or Nicaragua or Niger or Nigeria or "Northern Mariana Islands" or Nyasaland or Pakistan or Palau or Palestine or Panama or "Papua New Guinea" or Paraguay or Peru or Philipines or Philippines or Phillipines or Phillippines or Principe or Rhodesia or Romania or Roumania or Ruanda or Rumania or Russia or Russian or "Russian Federation" or Rwanda or Sabah or "Saint Lucia" or "Saint Vincent" or Samoa or "Samoan Islands" or "Sao Tome" or Sarawak or Senegal or Serbia or Seychelles or "Sierra Leone" or "Solomon Islands" or Somalia or "South Africa" or "South Sudan" or "Soviet Union" or "Sri Lanka" or "St Lucia" or "St Vincent" or Sudan or Surinam or Suriname or Swaziland or Syria or Syrian or Tadjikistan or Tadzhik or Tadzhikistan or Tajikistan or Tanzania or Thailand or "Timor Leste" or Togo or "Togolese Republic" or Tonga or Tunisia or Turkey or Turkmen or Turkmenistan or Tuvalu or Uganda or Ukraine or "Union of Soviet Socialist Republics" or "United Arab Republic" or "Upper Volta" or Uruguay or Urundi or USSR or Uzbek or Uzbekistan or Vanuatu or Venezuela or "Viet Nam" or Vietnam or "West Bank" or "Western Sahara" or Yemen or Yugoslavia or Zaire or Zambia or Zimbabwe).hw,kf,ti,ab,cp.

8. ((developing or less* developed or under developed or underdeveloped or middle income or low* income or underserved or under served or deprived or poor*) adj (countr* or nation? or population? or world or state*)).ti,ab.
9. ((developing or less* developed or under developed or underdeveloped or middle income or low* income) adj (economy or economies)).ti,ab.
10. (low* adj (gdp or gnp or gross domestic or gross national)).ti,ab.
11. (low adj3 middle adj3 countr*).ti,ab.
12. (lmic or lmics or third world or lami countr*).ti,ab.
13. transitional countr*.ti,ab.
14. (1 or 2) and or/3-13
15. Limit 14 to (chinese or english or french or spanish)

PAIS/ProQuest

1. su.explode("health" OR "public health")
2. su.exact("social networks" OR "social behaviour" OR "behavior modification" OR "externalities economics" OR "economic development" OR "economic development Social aspects")
3. ti,ab((“herd effect*1” AND (vaccin* OR immunis* OR immuniz*)) OR (cascade near/3 effect*1) or (indirect near/3 “causal effect*”) or (indirect* near/3 protect*) or

(infectiousness near/3 effect*) or spillover* or (spill*1 over) or externalit* or (interference and (vaccin* OR immunis* OR immuniz*)) OR (“vaccine efficacy” AND infectious*) OR (interference AND causal) or (peer* near/3 (effect* or influenc*)) or ((“social network*” or “social* interact*”) and (connect* or diffus* or spread* or (spill* adj over) or externalit* or contag* or infect* or transmis* or transmit*)) or “herd immunity” or “herd protect*” or “unexpected benefit*” or “unexpected effect*” or ((behavio* ad3 chang*or behavio* adj3 modif*) and (connect* or diffus* or spread* or (spill* adj over) or externalit* or contag* or infect* or transmis* or transmit*)))

4. la(english or spanish or chinese or french)
5. su.exact(“developing countries”)
6. ti,ab,su,loc(Africa or Asia or Caribbean or “West Indies” or “South America” or “Latin America” or “Central America”)((ti,ab((Afghanistan or "Agalega Islands" or Albania or Algeria or Angola or Antigua or Argentina or Armenia or Armenian or Azerbaijan or Bangladesh or Barbuda or Basutoland or Belarus or Belize or Belorussia or Belorussian or Benin or Bhutan or Bolivia or Bosnia or Botswana or Brasil or Brazil or Bulgaria or "Burkina Faso" or "Burkina Fasso" or Burma or Burundi or Byelarus or Byelorussian or Cambodia or Cameron or Camerons or Cameroon or Cameroons or "Cape Verde" or "Central African Republic" or Ceylon or Chad or Chile or China or Colombia or Comores or "Comoro Islands" or Comoros or Congo or "Costa Rica" or "Cote d'Ivoire" or Cuba or Djibouti or Dominica or "Dominican Republic" or "East Timor" or "East Timur" or Ecuador or Egypt or "El Salvador" or Eritrea or Ethiopia or Fiji or "French Somaliland" or Gabon or "Gabonese Republic" or Gambia or Gaza or "Georgia Republic" or "Georgian Republic" or Ghana or "Gold Coast" or Grenada or Grenadines or Guatemala or Guiana or Guinea or "Guinea Bissau" or Guyana or Haiti or Hercegovina or Herzegovina or Honduras or Ifni or India or "Indian Ocean" or Indonesia or Iran or Iraq or "Ivory Coast" or Jamahiriya or Jamahiryria or Jamaica or Jordan or Kampuchea or Kazakh or Kazakhstan or Kenya or "Khmer Republic" or Kirghiz or Kirghizia or Kirgizstan or Kiribati or Korea or Kosovo or "Kyrgyz Republic" or Kyrgyzstan or "Lao PDR" or Laos or Latvia or Lebanon or Lesotho or Liberia or Libia or Libya or Lithuania or Macedonia or Madagascar or "Malagasy Republic" or Malawi or Malay or Malaya or Malaysia or Maldives or Mali or "Marshall Islands" or Mauritania or Mauritius or Mayotte or Melanesia or Mexico or Micronesia or Mocambique or Moldova or Moldovia or Moldovian or Mongolia or Montenegro or Montenegro or Morocco or Mozambique or Myanma or Myanmar or Namibia or "Navigator Island" or "Navigator Islands" or Nepal or "New Caledonia" or "New Hebrides" or Nicaragua or Nicaragua or Niger or Nigeria or "Northern Mariana Islands" or Nyasaland or Pakistan or Palau or Palestine or Panama or "Papua New Guinea" or Paraguay or Peru or Philipines or Philippines or Phillipines or Phillippines or Principe or Rhodesia or Romania or Roumania or Ruanda or Rumania or Russia or Russian or "Russian Federation" or Rwanda or Sabah or "Saint Lucia" or "Saint Vincent" or Samoa or "Samoan Islands" or "Sao Tome" or Sarawak or Senegal or Serbia or Seychelles or "Sierra Leone" or "Solomon Islands" or Somalia or "South Africa" or "South Sudan" or "Soviet Union" or "Sri Lanka" or "St Lucia" or "St Vincent" or Sudan or Surinam or Suriname or Swaziland or Syria or Syrian or Tadjikistan or Tadzhik or Tadzhikistan or Tajikistan or Tanzania or Thailand or "Timor Leste" or Togo or "Togolese Republic" or Tonga or Tunisia or Turkey or Turkmen or Turkmenistan or Tuvalu or Uganda or Ukraine or "Union of Soviet Socialist Republics" or "United Arab Republic" or "Upper Volta" or Uruguay or Urundi or USSR

or Uzbek or Uzbekistan or Vanuatu or Venezuela or "Viet Nam" or Vietnam or "West Bank" or "Western Sahara" or Yemen or Yugoslavia or Zaire or Zambia or Zimbabwe))))

7. ti,ab((developing or "less* developed" or "under developed" or underdeveloped or "middle income" or "low* income" or underserved or "under served" or deprived or poor*) near (countr* or nation*1 or population*1 or world or state*))
8. ti,ab((developing or “less* developed” or “under developed” or underdeveloped or “middle income” or “low* income”) near (economy or economies))
9. ti,ab(low* near (gdp or gnp or “gross domestic” or “gross national”))
10. ti,ab(low near/3 middle near/3 countr*)
11. ti,ab(LMIC OR LMICS OR "third world" OR "LAMI countr*")
12. ti,ab("transitional countr*")
13. 1 and (2 or 3) and 4 and (5 or 6 or 7 or 8 or 9 or 10 or 11 or 12 or 13)

PsycInfo

1. ((((herd effect*1 and (vaccin* or immunis* or immuniz*)) or (cascade adj3 effect*1) or (indirect adj3 causal effect*) or indirect*) adj3 protect*) or (infectiousness adj3 effect*) or spillover* or spill*1 over or externalit* or (interference and (vaccin* or immunis* or immuniz*)) or (vaccine efficacy and infectious*) or (interference and causal) or (peer* adj3 (effect* or influenc*)) or ((social network* or social* interact*) and (connect* or diffus* or spread* or (spill* adj over) or externalit* or contag* or infect* or transmis* or transmit*)) or herd immunity or herd protect* or unexpected benefit* or unexpected effect* or ((behavio* adj3 chang*or behavio* adj3 modif*) and (connect* or diffus* or spread* or (spill* adj over) or externalit* or contag* or infect* or transmis* or transmit*))).ti,ab,hw.
2. interpersonal influences/ or interpersonal interaction/ or peer relations/ or social influences/ or social interaction/ or social networks/ or social behavior/ or social reinforcement/ or social facilitation/ or contagion/ or economic development/ or behavior change/ or behavior modification/
3. 1 or 2
4. (Africa or Asia or Caribbean or West Indies or South America or Latin America or Central America).ti,ab,hw.
5. (Afghanistan or "Agalega Islands" or Albania or Algeria or Angola or Antigua or Argentina or Armenia or Armenian or Azerbaijan or Bangladesh or Barbuda or Basutoland or Belarus or Belize or Belorussia or Belorussian or Benin or Bhutan or Bolivia or Bosnia or Botswana or Brasil or Brazil or Bulgaria or "Burkina Faso" or "Burkina Fasso" or Burma or Burundi or Byelarus or Byelorussian or Cambodia or Cameron or Camerons or Cameroon or Cameroons or "Cape Verde" or "Central African Republic" or Ceylon or Chad or Chile or China or Colombia or Comores or "Comoro Islands" or Comoros or Congo or "Costa Rica" or "Cote d'Ivoire" or Cuba or Djibouti or Dominica or "Dominican Republic" or "East Timor" or "East Timur" or Ecuador or Egypt or "El Salvador" or Eritrea or Ethiopia or Fiji or "French Somaliland" or Gabon or "Gabonese Republic" or Gambia or Gaza or "Georgia Republic" or "Georgian Republic" or Ghana or "Gold Coast" or Grenada or Grenadines or Guatemala or Guiana or Guinea or "Guinea Bissau" or Guyana or Haiti or Hercegovina or Herzegovina or Honduras or Ifni or India or "Indian Ocean" or Indonesia or Iran or Iraq or "Ivory Coast" or Jamahiriya or Jamahiryria or Jamaica or

Jordan or Kampuchea or Kazakh or Kazakhstan or Kenya or "Khmer Republic" or Kirghiz or Kirghizia or Kirgizstan or Kiribati or Korea or Kosovo or "Kyrgyz Republic" or Kyrgyzstan or "Lao PDR" or Laos or Latvia or Lebanon or Lesotho or Liberia or Libia or Libya or Lithuania or Macedonia or Madagascar or "Malagasy Republic" or Malawi or Malay or Malaya or Malaysia or Maldives or Mali or "Marshall Islands" or Mauritania or Mauritius or Mayotte or Melanesia or Mexico or Micronesia or Mocambique or Moldova or Moldovia or Moldovian or Mongolia or Montenegro or Montenegro or Morocco or Mozambique or Myanma or Myanmar or Namibia or "Navigator Island" or "Navigator Islands" or Nepal or "New Caledonia" or "New Hebrides" or Nicaragua or Nicaragua or Niger or Nigeria or "Northern Mariana Islands" or Nyasaland or Pakistan or Palau or Palestine or Panama or "Papua New Guinea" or Paraguay or Peru or Philipines or Philippines or Phillipines or Phillippines or Principe or Rhodesia or Romania or Roumania or Ruanda or Rumania or Russia or Russian or "Russian Federation" or Rwanda or Sabah or "Saint Lucia" or "Saint Vincent" or Samoa or "Samoan Islands" or "Sao Tome" or Sarawak or Senegal or Serbia or Seychelles or "Sierra Leone" or "Solomon Islands" or Somalia or "South Africa" or "South Sudan" or "Soviet Union" or "Sri Lanka" or "St Lucia" or "St Vincent" or Sudan or Surinam or Suriname or Swaziland or Syria or Syrian or Tadjikistan or Tadzhik or Tadzhikistan or Tajikistan or Tanzania or Thailand or "Timor Leste" or Togo or "Togolese Republic" or Tonga or Tunisia or Turkey or Turkmen or Turkmenistan or Tuvalu or Uganda or Ukraine or "Union of Soviet Socialist Republics" or "United Arab Republic" or "Upper Volta" or Uruguay or Urundi or USSR or Uzbek or Uzbekistan or Vanuatu or Venezuela or "Viet Nam" or Vietnam or "West Bank" or "Western Sahara" or Yemen or Yugoslavia or Zaire or Zambia or Zimbabwe).tw.

6. ((developing or less* developed or under developed or underdeveloped or middle income or low* income or underserved or under served or deprived or poor*) adj (countr* or nation? or population? or world or state*)).ti,ab.
7. ((developing or less* developed or under developed or underdeveloped or middle income or low* income) adj (economy or economies)).ti,ab.
8. (low* adj (gdp or gnp or gross domestic or gross national)).tw.
9. (low adj3 middle adj3 countr*).tw.
10. (lmic or lmics or third world or lami countr*).tw.
11. transitional countr*.tw.
12. or/4-11
13. 3 and 12
14. (chinese or english or french or spanish).lg.
15. 13 and 14

Web of Science

- Select Language=(English OR Chinese OR French OR Spanish)

1. TS=((“herd effect*” AND (vaccin* OR immunis* OR immuniz*)) OR (cascade near/3 effect*) or (indirect near/3 “causal effect*”) or indirect* near/3 protect* or (infectiousness near/3 effect*) or spillover* or (spill* over) or externalit* or (interference and (vaccin* OR immunis* OR immuniz*)) OR (“vaccine efficacy” AND infectious*) OR (interference AND causal) or (peer* near/3 (effect* or influenc*)) or ((“social network*” or “social* interact*”) and (connect* or diffus* or spread* or (spill*

adj over) or externalit* or contag* or infect* or transmis* or transmit*)) or “herd immunity” or “herd protect*” or “unexpected benefit*” or “unexpected effect*” or ((behavio* ad3 chang*or behavio* adj3 modif*) and (connect* or diffus* or spread* or (spill* adj over) or externalit* or contag* or infect* or transmis* or transmit*)))

2. TS=(Afghanistan or "Agalega Islands" or Albania or Algeria or Angola or Antigua or Argentina or Armenia or Armenian or Azerbaijan or Bangladesh or Barbuda or Basutoland or Belarus or Belize or Belorussia or Belorussian or Benin or Bhutan or Bolivia or Bosnia or Botswana or Brasil or Brazil or Bulgaria or "Burkina Faso" or "Burkina Fasso" or Burma or Burundi or Byelarus or Byelorussian or Cambodia or Cameron or Camerons or Cameroon or Cameroons or "Cape Verde" or "Central African Republic" or Ceylon or Chad or Chile or China or Colombia or Comores or "Comoro Islands" or Comoros or Congo or "Costa Rica" or "Cote d'Ivoire" or Cuba or Djibouti or Dominica or "Dominican Republic" or "East Timor" or "East Timur" or Ecuador or Egypt or "El Salvador" or Eritrea or Ethiopia or Fiji or "French Somaliland" or Gabon or "Gabonese Republic" or Gambia or Gaza or "Georgia Republic" or "Georgian Republic" or Ghana or "Gold Coast" or Grenada or Grenadines or Guatemala or Guiana or Guinea or "Guinea Bissau" or Guyana or Haiti or Hercegovina or Herzegovina or Honduras or Ifni or India or "Indian Ocean" or Indonesia or Iran or Iraq or "Ivory Coast" or Jamahiriya or Jamahiryria or Jamaica or Jordan or Kampuchea or Kazakh or Kazakhstan or Kenya or "Khmer Republic" or Kirghiz or Kirghizia or Kirgizstan or Kiribati or Korea or Kosovo or "Kyrgyz Republic" or Kyrgyzstan or "Lao PDR" or Laos or Latvia or Lebanon or Lesotho or Liberia or Libia or Libya or Lithuania or Macedonia or Madagascar or "Malagasy Republic" or Malawi or Malay or Malaya or Malaysia or Maldives or Mali or "Marshall Islands" or Mauritania or Mauritius or Mayotte or Melanesia or Mexico or Micronesia or Mocambique or Moldova or Moldovia or Moldovian or Mongolia or Montenegro or Montenegro or Morocco or Mozambique or Myanma or Myanmar or Namibia or "Navigator Island" or "Navigator Islands" or Nepal or "New Caledonia" or "New Hebrides" or Nicaragua or Nicaragua or Niger or Nigeria or "Northern Mariana Islands" or Nyasaland or Pakistan or Palau or Palestine or Panama or "Papua New Guinea" or Paraguay or Peru or Philipines or Philippines or Phillipines or Phillippines or Principe or Rhodesia or Romania or Roumania or Ruanda or Rumania or Russia or Russian or "Russian Federation" or Rwanda or Sabah or "Saint Lucia" or "Saint Vincent" or Samoa or "Samoan Islands" or "Sao Tome" or Sarawak or Senegal or Serbia or Seychelles or "Sierra Leone" or "Solomon Islands" or Somalia or "South Africa" or "South Sudan" or "Soviet Union" or "Sri Lanka" or "St Lucia" or "St Vincent" or Sudan or Surinam or Suriname or Swaziland or Syria or Syrian or Tadjikistan or Tadzhik or Tadzhikistan or Tajikistan or Tanzania or Thailand or "Timor Leste" or Togo or "Togolese Republic" or Tonga or Tunisia or Turkey or Turkmen or Turkmenistan or Tuvalu or Uganda or Ukraine or "Union of Soviet Socialist Republics" or "United Arab Republic" or "Upper Volta" or Uruguay or Urundi or USSR or Uzbek or Uzbekistan or Vanuatu or Venezuela or "Viet Nam" or Vietnam or "West Bank" or "Western Sahara" or Yemen or Yugoslavia or Zaire or Zambia or Zimbabwe)
3. TS=(Developing Countries) OR TS=(Africa or Asia or Caribbean or West Indies or South America or Latin America or Central America) OR TS=(((developing or "less* developed" or "under developed" or underdeveloped or "middle income" or "low*

income" or underserved or "under served" or deprived or poor*) NEAR/1 (countr* or nation* or population* or world)))

4. TS=(((developing or "less* developed" or "under developed" or underdeveloped or "middle income" or "low* income") NEAR/1 (economy or economies))) OR TS=((low* NEAR/1 (gdp or gnp or "gross domestic" or "gross national"))) OR TS=((low NEAR/3 middle NEAR/3 countr*))
5. TS=((lmic or lmics or "third world" or "lami countr*")) OR TS=(transitional countr*)
6. #1 AND (#2 OR #3 OR #4 OR #5)
7. Refined by: Web of Science Categories=(ECONOMICS OR PUBLIC ENVIRONMENTAL OCCUPATIONAL HEALTH OR PSYCHOLOGY MULTIDISCIPLINARY OR TROPICAL MEDICINE OR ANTHROPOLOGY OR DEMOGRAPHY OR PSYCHIATRY OR IMMUNOLOGY OR SOCIAL SCIENCES INTERDISCIPLINARY OR COMMUNICATION OR INFECTIOUS DISEASES OR HEALTH CARE SCIENCES SERVICES OR PSYCHOLOGY CLINICAL OR PEDIATRICS OR PSYCHOLOGY SOCIAL OR MULTIDISCIPLINARY SCIENCES OR NUTRITION DIETETICS OR SOCIAL SCIENCES BIOMEDICAL OR PARASITOLOGY OR RESPIRATORY SYSTEM OR PSYCHOLOGY OR PSYCHOLOGY APPLIED OR MEDICINE GENERAL INTERNAL OR BEHAVIORAL SCIENCES OR MEDICINE RESEARCH EXPERIMENTAL OR FAMILY STUDIES OR SOCIOLOGY OR AREA STUDIES OR HEALTH POLICY SERVICES OR URBAN STUDIES OR PSYCHOLOGY DEVELOPMENTAL OR EDUCATION EDUCATIONAL RESEARCH OR SOCIAL WORK)

WHO Global Health Library

- **Search all indexes (title, author, subject)**
- **Regional indexes only**
- **Download separate files for English, French, Spanish**

("herd effect*" AND (vaccin* OR immunis* OR immuniz*)) OR (cascade and effect*) OR (indirect AND "causal effect*") OR (indirect* AND protect*) OR (infectiousness AND effect*) OR spillover* OR "spill* over" OR externalit* OR (interference AND (vaccin* OR immunis* OR immuniz*)) OR ("vaccine efficacy" AND infectious*) OR (interference AND causal) OR (peer* AND (effect* OR influenc*)) OR ("social network*" AND (connect* OR diffus* OR spread* OR (spill* AND over) OR externalit* OR contag* OR infect* OR transmis* OR transmit*)) OR "herd immunity" OR "herd protect*" OR "unexpected benefit*" OR "unexpected effect*" OR (((behavio* AND chang*) OR (behavio* AND modif*)) AND (connect* OR diffus* OR spread* OR (spill* AND over*) OR spillover* OR externalit* OR contag* OR infect* OR infectious* OR transmis* OR transmit*))

WPSA/ProQuest

1. su.explode("Community Mental Health" OR "Health" OR "Mental Health" OR "Occupational Safety and Health" OR "Public Health" OR "Economic development")
2. su.exact.explode("Health Behavior" OR "Health Care Services" OR "Health Maintenance Organizations" OR "Home Health Care" OR "Long Term Care" OR "Managed Care Services" OR "Mental Health Services" OR "Primary Health Care" OR "Womens Health Care" OR "Health Education")
3. SU.EXACT.EXPLODE("Behavior Modification" OR "Social Influence")

4. ti,ab((“herd effect*1” AND (vaccin* OR immunis* OR immuniz*)) OR (cascade near/3 effect*1) or (indirect near/3 “causal effect*”) or (indirect* near/3 protect*) or (infectiousness near/3 effect*) or spillover* or (spill*1 over) or externalit* or (interference and (vaccin* OR immunis* OR immuniz*)) OR (“vaccine efficacy” AND infectious*) OR (interference AND causal) or (peer* near/3 (effect* or influenc*)) or ((“social network*” or “social* interact*”) and (connect* or diffus* or spread* or (spill* adj over) or externalit* or contag* or infect* or transmis* or transmit*)) or “herd immunity” or “herd protect*” or “unexpected benefit*” or “unexpected effect*” or ((behavio* ad3 chang*or behavio* adj3 modif*) and (connect* or diffus* or spread* or (spill* adj over) or externalit* or contag* or infect* or transmis* or transmit*)))
5. la(english or spanish or chinese or french)
6. su.exact(“developing countries”)
7. ti,ab,su,loc(Africa or Asia or Caribbean or “West Indies” or “South America” or “Latin America” or “Central America”)
8. ((ti,ab((Afghanistan or "Agalega Islands" or Albania or Algeria or Angola or Antigua or Argentina or Armenia or Armenian or Azerbaijan or Bangladesh or Barbuda or Basutoland or Belarus or Belize or Belorussia or Belorussian or Benin or Bhutan or Bolivia or Bosnia or Botswana or Brasil or Brazil or Bulgaria or "Burkina Faso" or "Burkina Fasso" or Burma or Burundi or Byelarus or Byelorussian or Cambodia or Cameron or Camerons or Cameroon or Cameroons or "Cape Verde" or "Central African Republic" or Ceylon or Chad or Chile or China or Colombia or Comores or "Comoro Islands" or Comoros or Congo or "Costa Rica" or "Cote d'Ivoire" or Cuba or Djibouti or Dominica or "Dominican Republic" or "East Timor" or "East Timur" or Ecuador or Egypt or "El Salvador" or Eritrea or Ethiopia or Fiji or "French Somaliland" or Gabon or "Gabonese Republic" or Gambia or Gaza or "Georgia Republic" or "Georgian Republic" or Ghana or "Gold Coast" or Grenada or Grenadines or Guatemala or Guiana or Guinea or "Guinea Bissau" or Guyana or Haiti or Hercegovina or Herzegovina or Honduras or Ifni or India or "Indian Ocean" or Indonesia or Iran or Iraq or "Ivory Coast" or Jamahiriya or Jamahiryria or Jamaica or Jordan or Kampuchea or Kazakh or Kazakhstan or Kenya or "Khmer Republic" or Kirghiz or Kirghizia or Kirgizstan or Kiribati or Korea or Kosovo or "Kyrgyz Republic" or Kyrgyzstan or "Lao PDR" or Laos or Latvia or Lebanon or Lesotho or Liberia or Libia or Libya or Lithuania or Macedonia or Madagascar or "Malagasy Republic" or Malawi or Malay or Malaya or Malaysia or Maldives or Mali or "Marshall Islands" or Mauritania or Mauritius or Mayotte or Melanesia or Mexico or Micronesia or Mocambique or Moldova or Moldovia or Moldovian or Mongolia or Montenegro or Montenegro or Morocco or Mozambique or Myanma or Myanmar or Namibia or "Navigator Island" or "Navigator Islands" or Nepal or "New Caledonia" or "New Hebrides" or Nicaragua or Nicaragua or Niger or Nigeria or "Northern Mariana Islands" or Nyasaland or Pakistan or Palau or Palestine or Panama or "Papua New Guinea" or Paraguay or Peru or Philipines or Philippines or Phillipines or Phillippines or Principe or Rhodesia or Romania or Roumania or Ruanda or Rumania or Russia or Russian or "Russian Federation" or Rwanda or Sabah or "Saint Lucia" or "Saint Vincent" or Samoa or "Samoan Islands" or "Sao Tome" or Sarawak or Senegal or Serbia or Seychelles or "Sierra Leone" or "Solomon Islands" or Somalia or "South Africa" or "South Sudan" or "Soviet Union" or "Sri Lanka" or "St Lucia" or "St Vincent" or Sudan or Surinam or Suriname or Swaziland or Syria or Syrian or Tadjikistan or Tadzhik or Tadzhikistan or Tajikistan or Tanzania or Thailand or "Timor Leste" or

Togo or "Togolese Republic" or Tonga or Tunisia or Turkey or Turkmen or Turkmenistan or Tuvalu or Uganda or Ukraine or "Union of Soviet Socialist Republics" or "United Arab Republic" or "Upper Volta" or Uruguay or Urundi or USSR or Uzbek or Uzbekistan or Vanuatu or Venezuela or "Viet Nam" or Vietnam or "West Bank" or "Western Sahara" or Yemen or Yugoslavia or Zaire or Zambia or Zimbabwe))))

9. ti,ab((developing or "less* developed" or "under developed" or underdeveloped or "middle income" or "low* income" or underserved or "under served" or deprived or poor*) near (countr* or nation*1 or population*1 or world or state*))
10. ti,ab((developing or “less* developed” or “under developed” or underdeveloped or “middle income” or “low* income”) near (economy or economies))
11. ti,ab(low* near (gdp or gnp or “gross domestic” or “gross national”))
12. ti,ab(low near/3 middle near/3 countr*)
13. ti,ab(LMIC OR LMICS OR "third world" OR "LAMI countr*")
14. ti,ab("transitional countr*")
15. (1 or 2) and (3 or 4) and 5 and (6 or 7 or 8 or 9 or 10 or 11 or 12 or 13 or 14)

<u>Non-bibliographic search strategy</u>

3ie Impact Evaluation Database

- **Under “Advanced” search “All Evidence”**
- **Make sure to clear searches in between each new search**

1. herd effect
2. herd immunity
3. herd protection
4. cascade effect
5. contagion
6. contagious
7. diffuse
8. indirect causal effect
9. indirect protection
10. infectiousness effect
11. spillover
12. spill over
13. externality
14. externalities
15. interference
16. vaccine efficacy
17. peer effect
18. peer influence
19. social network
20. social interaction
21. unexpected benefit
22. unexpected effect
23. behavior change
24. behavioral change
25. behavior modification
26. behavioral modification

27. modify behavior
28. modifies behavior
29. behaviour change
30. behavioural change
31. behaviour modification
32. behavioural modification
33. modify behaviour
34. modifies behaviour

Campbell Library

- **Search all text**

1. “herd effect”
2. “herd effects”
3. “herd immunity”
4. “herd protection”
5. “cascade effect”
6. “cascade effects”
7. diffuse
8. diffusion
9. “indirect causal effect”
10. “indirect causal effects”
11. “indirect protection”
12. “infectiousness effect”
13. “infectiousness effects”
14. interference
15. vaccine efficacy
16. spillover
17. spillovers
18. “spill over”
19. “spills over”
20. “spill overs”
21. externality
22. externalities
23. “peer effect”
24. “peer effects”
25. “peers’ effects”
26. “peer influence”
27. “peer influences”
28. “peers’ influences”
29. “social network”
30. “social networks”
31. “social interaction”
32. “social interactions”
33. “unexpected benefit”
34. “unexpected benefits”
35. “unexpected effect”
36. “unexpected effects”

37. “behavior change”
38. “behavioral change”
39. “behavior changes”
40. “behavior modification”
41. “behavior modifications”
42. “behavioral modifications”
43. “modify behavior”
44. “modify behaviors”
45. “modifies behavior”
46. “modifies behaviors”
47. “behaviour change”
48. “behavioural change”
49. “behaviour changes”
50. “behaviour modification”
51. “behaviour modifications”
52. “behavioural modifications”
53. “modify behaviour”
54. “modify behaviours”
55. “modifies behaviour”
56. “modifies behaviours”

Center for Global Development

- **Title only search**

1. herd effect
2. herd immunity
3. herd protection
4. cascade effect
5. contagion
6. contagious
7. diffuse
8. indirect causal effect
9. indirect protection
10. infectiousness effect
11. spillover
12. spill over
13. spills over
14. externality
15. externalities
16. interference
17. vaccine efficacy
18. peer effect
19. peer influence
20. social network
21. social interaction
22. unexpected benefit
23. unexpected effect
24. behavior change

25. behavioral change
26. behavior modification
27. behavioral modification
28. modify behavior
29. modifies behavior
30. behavior change
31. behavioural change
32. behaviour modification
33. behavioural modification
34. modify behaviour
35. modifies behaviour

Center for Reviews and Dissemination

- **Search in "Any field"**
- **Separate terms with Boolean operators by entering them into separate boxes**

1. herd effect* AND (vaccin* OR immunis* OR immuniz*)
2. herd immunity OR herd protect*
3. cascade effect*
4. indirect causal effect* OR indirect* protect*
5. infectiousness effect*
6. spillover* OR spill* over*
7. externalit*
8. interference AND causal
9. interference AND (vaccin* OR immunis* OR immuniz*)
10. vaccine efficacy AND infectious*
11. peer* effect* OR peer* influenc*
12. social network* AND (connect* or diffus* or spread* or spillover* or externalit* or contag* or infect* or transmis* or transmit*)
13. social interaction* AND (connect* or diffus* or spread* or spillover* or externalit* or contag* or infect* or transmis* or transmit*)
14. unexpected benefit* OR unexpected effect*
15. behav* chang* AND (connect* or diffus* or spread* or spillover* or externalit* or contag* or infect* or transmis* or transmit*)
16. behav* modif* AND (connect* or diffus* or spread* or spillover* or externalit* or contag* or infect* or transmis* or transmit*)
17. modif* behav* AND (connect* or diffus* or spread* or spillover* or externalit* or contag* or infect* or transmis* or transmit*)

DFID R4D Database

- **Go to "Advanced Search"**
- **Select "Find Documents"**
- **Choose "All these fields" and "Exact phrase" unless otherwise specified**
- **Use drop down boxes to input Boolean operators**

1. herd effect* OR herd immunity OR herd protection
2. cascade effect*
3. contag*
4. diffus*

5. indirect causal effect* OR indirect protection
6. infectiousness effect*
7. spillover* OR spill over* OR spills over
8. externalit*
9. **All these words:** interference AND (vaccin* immuniz* immunis*)
10. Interference AND causal
11. vaccine efficacy
12. peer effect* OR peer influence*
13. peers' effect* OR peers' influence*
14. social network* OR social interact*
15. unexpected benefit* OR unexpected effect*
16. behavior chang* OR behavioral chang*
17. behaviour chang* OR behavioural chang*
18. behavior modif* OR behavioral modif*
19. behaviour modif* OR behavioural modif*
20. modify behavior* OR modify behaviour*

Google/Google Scholar

1. herd effect vaccin immunis immuniz immunity
2. "indirect causal effect" OR "indirect protection" OR "infectiousness effect" health
3. interference vaccin immunis immuniz immunity
4. peer effects influence health
5. social network interaction health
6. "unexpected benefit" OR "unexpected effect" health
7. behav chang modif health
8. spillover health
9. externalities health economic development

IDEAS

- **Choose Match "Boolean", in "Abstract"**
- **Restrict to "articles, papers, chapters, books"**

1. "herd effect"
2. "herd immunity"
3. "herd protection"
4. "cascade effect"
5. "indirect causal effect"
6. "indirect protection"
7. "infectiousness effect"
8. spillover
9. "spill over"
10. externality
11. externalities
12. "vaccine efficacy"
13. "social network"
14. "social interaction"
15. causal interference
16. "peer effect"

17. “peer influence”
18. “behavior change”
19. “behaviour change”
20. “behavior modification”
21. “behaviour modification”
22. “modify behavior”
23. “modify behaviour”

JOLIS

- **In Advanced Search, Choose Title**
- **Copy and paste items between Boolean operators into separate boxes in the search window**

1. herd effect OR herd protection OR herd immunity
2. cascade effect
3. indirect causal effect OR indirect protection
4. infectiousness effect
5. spillover OR spill over
6. externality OR externalities
7. interference AND vaccine
8. interference AND immunize
9. interference AND immunize
10. interference AND causal
11. vaccine efficacy
12. peer effect OR peer influence
13. social network OR social interaction
14. unexpected effect OR unexpected benefit
15. behavior change OR behaviour change OR behavior modification OR behaviour modification

NBER

- **Choose author/title search, working papers, books**

1. “herd effect”
2. “herd effects”
3. “herd immunity”
4. “herd protection”
5. “cascade effect”
6. “cascade effects”
7. “indirect causal effect”
8. “indirect causal effects”
9. “indirect protection”
10. “infectiousness effect”
11. “infectiousness effects”
12. spillover
13. spillovers
14. “spill over”
15. “spills over”
16. externality

17. externalities
18. “vaccine efficacy”
19. “social network”
20. “social networks”
21. “social interaction”
22. “social interactions”
23. “causal interference”
24. “peer effect”
25. “peer effects”
26. “peer influence”
27. “peer influences”
28. “behavior change”
29. “behaviour change”
30. “behavior modification”
31. “behaviour modification”
32. “modify behavior”
33. “modifies behavior”
34. “modify behaviour”
35. “modifies behaviour”

OpenGrey

- **Choose full text search of publications, working papers, books**

1. Abstract: "herd effect" OR "herd effects" OR "herd immunity" OR "herd protection" OR "cascade effect" OR "cascade effects" OR "indirect causal effect" OR "indirect causal effects" OR "indirect protection" OR "infectiousness effect" OR "infectiousness effects" OR spillover* OR "spill over" OR "spills over" OR externalit* OR "vaccine efficacy" OR "social network" OR "social networks" OR "social interaction" OR "social interactions" OR "causal interference" OR "peer effect" OR "peer effects" OR "peer influence" OR "peer influences" OR "behavior change" OR "behaviour change" OR "behavior modification" OR "behaviour modification" OR "modify behavior" OR "modifies behavior" OR "modify behaviour" OR "modifies behaviour"

SSRN

- **Search title, abstract, abstract ID, keywords, all dates**

1. “herd effect”
2. “herd effects”
3. “herd immunity”
4. “herd protection”
5. “cascade effect”
6. “cascade effects”
7. “indirect causal effect”
8. “indirect causal effects”
9. “indirect protection”
10. “infectiousness effect”
11. “infectiousness effects”
12. spillover
13. spillovers

14. “spill over”
15. “spills over”
16. externality
17. externalities
18. “vaccine efficacy”
19. “social network”
20. “social networks”
21. “social interaction”
22. “social interactions”
23. “causal interference”
24. “peer effect”
25. “peer effects”
26. “peer influence”
27. “peer influences”
28. “behavior change”
29. “behaviour change”
30. “behavior modification”
31. “behaviour modification”
32. “modify behavior”
33. “modifies behavior”
34. “modify behaviour”

Appendix 2. Electronic databases searched

“modifies behaviour” in Global Evaluation Database
The Cochrane Library
Econlit
EMBASE
ERIC
Google (first 300 results sorted by relevance)
Google Scholar (first 300 results sorted by relevance)
International Bibliography of Social Sciences (IBSS)
IDEAS
JOLIS library catalogue – International Monetary Fund
NBER Working Papers
PsycINFO
PubMed (includes MEDLINE)
Public Affairs Info Service (PAIS)
Social Science Research Network
Web of Science
WHO Global Health Library
Worldwide Political Science Abstracts (WPSA)

* We searched the following databases as well but did not review titles from these databases due to the large volume of results returned, and in some cases, the limited ability to control search queires in these databases: The Campbell Library, Center for Global Development Publications, Center for Reviews and Dissemination, Chinese National Knowledge Infrastructure Database, DFID R4D Database, OpenGrey, Wan Fang

Appendix 3: Risk of bias assessment

Internal validity	**% Yes**	**% Uncertain**	**% No**	**No. applicable papers**
Intervention and control groups have similar characteristics at baseline [f]	76%	20%	4%	49
Outcome data was collected at the same time and using the same methods for the intervention and control groups [a]	98%	0%	2%	51
If outcome data was reported to be missing, reasons for missing data are acceptable (e.g. missingness is non-differential between treatment groups) [b]	63%	29%	8%	24
Whether there was missing outcome data was reported	48%	0%	52%	48
For the estimation of direct effects, there was evidence of minimal contamination of the control group (not assessed for spillover effect estimation) [a]	19%	49%	32%	37
Subjects included in estimation of direct effects are comparable to subjects included in estimation of spillover effects	86%	2%	12%	42
Authors describe the extent to which the comparison group used to estimate spillovers could be considered a valid counterfactual	89%	4%	7%	54
Randomized designs only				
Random assignment at the appropriate level; for instance, an intervention that is delivered at the group level is randomized at the group level. [a]	100%	0%	0%	35
Where appropriate, study investigators and participants were blinded to treatment assignment [b]	34%	11%	55%	38
The allocation sequence used to assign treatment randomly was adequately generated [b]	49%	37%	14%	35
The allocation was adequately concealed [b]	46%	50%	4%	28
Study investigators were blinded to outcome assessment. [b]	32%	58%	11%	38
Case-control studies only				
Factors considered to be strong confounders are controlled for [c]	33%	0%	67%	3
Controls are selected from the same population as the cases [c]	100%	0%	0%	3
In studies examining the first occurrence among cases, controls with previous occurrences are excluded [c]	50%	0%	50%	2

Exposure is ascertained through secure records or structured interviews which are blinded to case/control status [c]	33%	67%	0%	3
Exposure is ascertained in the same way for cases and controls [c]	100%	0%	0%	3
Cohort and difference-in-difference studies only				
Treatment status was ascertained through secure records or structured interviews [c]	100%	0%	0%	10
The outcome was not present in the population at the beginning of the study [c]	75%	0%	0%	4
Cohort studies only				
Factors considered to be strong confounders are controlled for [c]	69%	0%	31%	13
Difference-in-difference studies only				
Equal trends assumption was checked: trends in the outcome are the same in both the treatment and control group prior to treatment [d]	20%	20%	60%	5
Validity of the equal trends assumption was explored (e.g. using a placebo test, “fake” treatment group, or “fake” control group) [d]	0%	20%	80%	5
Regression discontinuity / Interrupted time series studies only				
Researchers investigated whether study participants attempted to manipulate their value of the intervention assignment variable in order to determine group assignment [e]	100%	0%	0%	2
Noncompliance around the cutoff for treatment was assessed [d]	100%	0%	0%	2
There is a strong discontinuity around the cutoff (i.e. is there a strong first-stage)	100%	0%	0%	2
Other variables change discontinuously around the cutoff	100%	0%	0%	2
Matched studies only				
Baseline characteristics for the intervention group and potential controls were measured at the same time and in the same way [d]	80%	0%	20%	5
Observed characteristics between matched groups are well balanced [d]	40%	40%	20%	5
Common support in the distribution of variables or propensity scores used to match was assessed and found to be sufficient [d]	60%	40%	0%	5
Instrumental variables only				
The exclusion restriction does not suffer from	0%	67%	33%	3

any obvious violations				
There is evidence that the instrument used is "strong"	33%	33%	33%	3
Regression discontinuity / Interrupted time series studies only				
The index used to determine eligibility was continuous around the cutoff score at baseline [d]	0%	100%	0%	1
Case-control studies only				
Case definition is adequate [c]	75%	0%	0%	4
Other				
The study reports estimated effect measures, accompanying standard error and/or 95% confidence interval, and number of units used to estimate it [a]	79%	0%	19%	53
The study reports effect measures for all primary outcomes measured rather than only for those with desired results. Evaluating this potential source of bias requires that authors either list all outcomes they measured and could have reported or that the study's protocol is registered. [b]	13%	85%	2%	54
Regression analysis only				
The observed effect is robust to a variety of specifications (e.g. inclusion of potential relevant controls/covariates)	76%	24%	0%	21
Standard errors are appropriately clustered if necessary	83%	14%	2%	42
If the study uses fixed effects to control for unobservable factors across time, individuals, etc., it considers non-fixed factors that may vary with the intervention	40%	0%	60%	5

[a] From the Coalition for Evidence-Based Policy's Checklist for Reviewing a Randomized Controlled Trial of a Social Program or Project (Coalition for Evidence-Based Policy 2010)

[b] Cochrane Handbook. (Higgins and Greene 2011)

[c] Detailed criteria will be based on the Newcastle-Ottawa Scale (Wells et al. 2011)

[d] (Gertler 2010)

[e] (Lee and Lemieux 2010)

[f] From the Effective Practice and Organisation of Care (EPOC) Group's Risk of Bias Tool (Effective Practice and Organisation of Care (EPOC) Group 2009)

Appendix 4. Rationale for the classification of each type of study design for estimating spillover effects as low, moderate, or high quality.

Underlying methodology with high quality evidence

- **Double-randomized trials estimating within-cluster spillovers:** This design allows for the best inference for within-cluster spillovers by minimizing selection bias within clusters and unmeasured confounding.
- **Cluster-randomized trials estimating within-cluster spillovers among people who were not eligible but were highly comparable to eligible individuals:** Investigators can utilize matching algorithms to maximize comparability of untreated individuals in the treatment arm to individuals in the control arm to minimize confounding. As long as ineligible individuals are comparable to eligible individuals (i.e., treatment group assignment is ignorable), there is no reason to believe the intervention would have different effects in ineligible individuals, and there is a large enough number of ineligible individuals for sufficient statistical power, this design should yield nearly the same quality of evidence as a double-randomized design.
- **Individually randomized studies estimating spillover effects among social network members:** This study can also minimize bias and produce valid measures of social network spillovers as long as the network of individuals measured for each randomized individual is not connected to other randomized individuals. Thus, such designs will produce the best quality evidence when enrolling a relatively large, diffuse study population.
- **Studies estimating spillovers conditional on treatment or outcome density in a randomized study in which treatment or outcome density is estimated over multiple distances (physical or social) and in which distance cutoffs are defined based on quantiles or other objective criteria:** Because randomization should ensure that the distribution of individuals allocated to treatment is not associated with the outcome or potential confounding factors, this design can allow for rigorous estimates of spillovers. It is preferable for studies to estimate spillovers conditional on density at different distances to allow for assessment of a possible dose-response pattern. In addition, to avoid bias associated with definition of areas in which density is measured (Openshaw 1984), studies that objectively pre-specify distance levels (e.g., using quantiles of the observed distance distribution) will produce the best quality evidence.
- **Household-based studies estimating vaccine efficacy parameters which match index cases with household controls:** This study design uses an observational study design and would not be classified as high quality under the Cochrane GRADE criteria. However, we consider this a rigorous study design for estimating vaccine efficacy parameters for certain interventions and outcomes. For example, in studies measuring vaccine-preventable outcomes, such as cholera or pertussis, factors that may affect a case or control's choice to receive a vaccine are unlikely to be strongly associated with the likelihood of contracting the illness. Thus, confounding is likely to be minimal in the observational design for vaccine efficacy parameters. These studies will produce the best quality evidence when households are located far apart in order to minimize spillovers between households.

Underlying methodology with moderate quality evidence

- **Randomized trials estimating within-cluster spillovers among people who chose not to participate in the intervention (i.e., participants within clusters were not randomized to receive treatment, so selection bias is possible in spillover effects)**: Because individuals who choose to participate in an intervention may be systematically different from those who choose not to, this study design produces lower quality evidence than designs that utilize cluster-randomized trials and measure spillovers among ineligible individuals.
- **Cluster-randomized trials estimating within-cluster spillovers among people who were not eligible and were not highly comparable to eligible individuals:** Eligibility criteria affect the extent to which spillover and direct effects are comparable. For many outcomes, such as trachoma, interventions are not likely to have substantially different effects based on age; in other words, age is not a confounder. However, income-based eligibility criteria are common in economic programs targeting the poor, such as conditional cash transfer programs. For such interventions, one would expect that the effect of the intervention could differ substantially between eligible and non-eligible individuals. Thus, we considered such studies to have moderate quality evidence.
- **Observational studies estimating within-cluster spillovers:** We consider these studies to have moderate quality evidence because they are unable to minimize unmeasured confounding.
- **Studies estimating spillovers conditional on treatment or outcome density in a randomized study in which treatment or outcome density is estimated over only one distance level (physical or social):** If a study only estimates spillovers conditional on treatment or outcome density within one distance level, spillover estimates are likely to be sensitive to the choice of distance or area definition (Openshaw 1984), particularly if the choice is neither grounded in biological or social theory nor pre-specified.
- **Studies estimating spillovers conditional on treatment or outcome density in a randomized study in which treatment or outcome density is estimated and distance cutoffs were not based on objective criteria:** If the distance or area within which density is measured is not defined objectively, spillover estimates are also likely to be sensitive to the choice of distance or area definition (Openshaw 1984), particularly if the choice is neither grounded in biological or social theory nor pre-specified.
- **Ecologic studies comparing outcomes over levels of treatment coverage in which the treatment was randomized and a possible dose-response pattern for spillovers was assessed:** For spillovers of health outcomes, which is the focus of this review, in nearly all cases, the desired level of inference is at the individual rather than group level. Thus ecologic studies will never produce as strong of evidence as studies that assess individual-level outcomes because the associations at the group-level cannot be assumed hold true at the individual level. In ecologic studies that measure the association between treatment coverage and outcomes, if treatment was randomized, the evidence for spillovers can still be of moderate quality because randomization can ensure that the distribution of individuals allocated to treatment is not associated with the outcome or potential confounding factors.

Underlying methodology with low or very low quality evidence

- **Ecological studies comparing outcomes over levels of treatment coverage in which the treatment was not randomized:** For reasons discussed above, ecological studies produce lower quality evidence of spillovers. Those that assess associations with treatment coverage in studies in which treatment was not randomized produce low quality evidence because the distribution of treatment may be associated with the outcome or potential confounding factors.
- **Ecological studies comparing outcomes over levels of treatment coverage that did not assess a possible dose-response gradient for spillover effects:** For reasons discussed above, ecological studies produce lower quality evidence of spillovers. Those that assess associations with treatment coverage in studies but only consider one or two levels of treatment coverage produce low quality evidence because a dose-response pattern cannot be assessed and because findings are sensitive to the choice of coverage levels that are compared.
- **Studies in which instrumental variables were the primary identification strategy but the exclusion restriction suffers from obvious violations or the instrument is not strongly associated with the treatment:** In order for an instrumental variable to be used to identify a causal effect in a study in which treatment was not randomized, three criteria must be met: 1) the instrument must be independent from confounders of the treatment and outcome, 2) the instrument must be associated with the treatment, and 3) the instrument must be independent of the outcome given the treatment and confounders of the treatment and outcome (Greenland 2000). If any of these criteria are not met, the quality of evidence produced by an instrumental variables analysis will be low.
- **Studies in which systematic differences were likely to be present between intervention and control group (e.g., a cohort study that did not using matching to make the control group comparable to the intervention group):** In such studies, spillover estimates are likely to be biased either towards or away from the null, thus the quality of evidence is low.
- **Studies that did not include a rigorous control group:** Such studies cannot adequately control for confounding and produce low evidence.

Appendix 5. Comprehensive validity assessment

External validity	% Yes	% Uncertain	% No	No. applicable papers
For prospective studies, participants were asked to give informed consent prior to assignment to a study group [a]	54%	46%	0%	54
Percent of people invited who consented to participate in the study is over 90%. This refers to participation in the activities of the impact evaluation rather than participation in the intervention itself.	11%	73%	16%	44
The authors comment about the representativeness of their sample and/or external validity	80%	6%	15%	54
Cohort and difference-in-difference studies only				
The treated cohort is representative of the target population [c]	69%	23%	8%	13
The non-treated cohort was drawn from the same population as the treated cohort [c]	77%	8%	15%	13
Case-control studies only				
Cases are representative [c]	67%	33%	0%	3
Construct validity	**% Yes**	**% Uncertain**	**% No**	**No. applicable papers**
The study's outcome measures are likely to be highly correlated with the true outcomes of interest [a]	93%	2%	6%	54
If intermediate outcomes are measured, their connection to terminal outcomes of interest is clearly defined [a]	90%	0%	10%	21
The study period was long enough to see meaningful changes in outcome measures [a]	96%	4%	0%	52
The authors discuss the possibility of error in the measurement of their outcome, treatment, and covariates	44%	0%	56%	54

Appendix 6. Reasons for exclusion of full texts

Reason for exclusion*	**Count****
No spillovers measured	429
Design not valid	51
No health outcomes	36
Not an empirical study	33
Not conducted in low-middle income country	30
No intervention evaluated	26
Wrong record type (e.g. conference abstract)	24
Review	6
Quantitative measures of spillover not presented	4
Outcomes not measured in humans	3
Qualitative study	2
Methods of estimating spillovers not described	2
Wrong intervention type	1
Total	647

*Only one reason was recorded for each record. In some cases, multiple reasons were applicable but only one was recorded.

**Includes full texts from original search, snowball search, and reference list searches.

Appendix 7. Overall quality of evidence for each included study

Authors	Quality rating
Ali et al., 2005	moderate
Ali et al., 2008	moderate
Ali et al., 2013	moderate
Avitabile, 2012	low
Azad et al., 2010	low
Baird et al., 2013	very low
Baird et al., 2013	low
Banerjee et al., 2010	high
Baptista et al., 2006	moderate
Bhattacharya et al. , 2013	low
Bjorkman et al., 2009	moderate
Buttenheim, 2011	low
Chaudhuri (Year not listed)	very low
Chen et al., 2014	low
Chidambaram et al., 2004	low
Chong, 2013	moderate
Contreras & Maitra, 2013	very low
Cooper & Fitch, 1983	very low
Dupas, 2006	moderate
Egere et al., 2012	high
Emch et al., 2006	moderate
Emch et al., 2009	moderate
Fitzsimons, 2012	moderate
Forleo-Neto et al., 1999	low
German, 2012	moderate
Godlonton & Thornton, 2012	moderate
Haile et al., 2013	very low
Hammitt et al., 2014	moderate
Handa, 2001	moderate
Hawley et al., 2003	high
House et al., 2009	high
Huq et al., 2010	moderate
Janssens, 2005	low
Joshi, 2013	very low
Kazianga et al., 2014	moderate
Khan et al., 2012	moderate
Khan et al., 2012	moderate

Khatib et al., 2012	low
Miguel & Kremer, 2003	moderate
Miguel & Kremer, 2004	moderate
Ozier, 2011	moderate
Paul, 1962	low
Perez-Heydrich et al., 2014	high
Preziosi et al., 2003	high
Ribas, 2011	moderate
Roca et al., 2011	moderate
Roca et al., 2013	moderate
Root et al., 2011	moderate
Root, 2014	moderate
Shekhawat, 2014	low
Singh (Year not listed)	moderate
Sur et al., 2009	moderate
Tontarawongsa, 2011	moderate
Ziegelhofer, 2012	moderate
Zivin, 2009	moderate

International Initiative for Impact Evaluation
London International Development Centre
36 Gordon Square
London WC1H 0PD
United Kingdom

3ieuk@3ieimpact.org
Tel: +44 207 958 8351/8350

www.3ieimpact.org