



Published in final edited form as:

AIDS Behav. 2011 August ; 15(6): 1075–1087. doi:10.1007/s10461-010-9847-0.

HIV Interventions to Reduce HIV/AIDS Stigma: A Systematic Review

Sohini Sengupta,

Center For Faculty Excellence, The University of North Carolina at Chapel Hill (UNC-Chapel Hill), Wilson Library, CB#3470, Chapel Hill, NC 27599-3470, USA

Bahby Banks,

Cecil G. Sheps Center for Health Services Research, UNC-Chapel Hill, Chapel Hill, NC, USA

Dan Jonas,

Department of Medicine, School of Medicine, UNC-Chapel Hill, Chapel Hill, NC, USA

Margaret Shandor Miles, and

School of Nursing, UNC-Chapel Hill, Chapel Hill, NC, USA

Giselle Corbie Smith

Department of Social Medicine, School of Medicine, UNC-Chapel Hill, Chapel Hill, NC, USA.
Cecil G. Sheps Center for Health Services Research, UNC-Chapel Hill, Chapel Hill, NC, USA.
Department of Medicine, School of Medicine, UNC-Chapel Hill, Chapel Hill, NC, USA. Division of General Medicine and Clinical Epidemiology, UNC-Chapel Hill, Chapel Hill, NC, USA.
Department of Epidemiology, School of Public Health, UNC-Chapel Hill, Chapel Hill, NC, USA

Sohini Sengupta: sengups@unc.edu

Abstract

We reviewed the literature to determine the effectiveness of HIV-related interventions in reducing HIV/AIDS stigma. Studies selected had *randomized controlled trial (RCT)*, *pretest–posttest with a non-randomized control group*, or *pretest–posttest one group* study designs in which HIV-related interventions were being evaluated, and in which HIV/AIDS stigma was one of the outcomes being measured. A checklist was used to extract data from accepted studies, assess their internal validity, and overall quality. Data were extracted from 19 studies, and 14 of these studies demonstrated effectiveness in reducing HIV/AIDS stigma. Only 2 of these 14 effective studies were considered good studies, based on quality, the extent to which the intervention focused on reducing HIV/AIDS stigma, and the statistics reported to demonstrate effectiveness. Future studies to reduce HIV/AIDS stigma could improve by designing interventions that pay greater attention to internal validity, use validated HIV/AIDS stigma instruments, and achieve both statistical and public health significance.

Keywords

HIV; AIDS; Review; Intervention studies; Program evaluation

Correspondence to: Sohini Sengupta, sengups@unc.edu.

Electronic supplementary material The online version of this article (doi:10.1007/s10461-010-9847-0) contains supplementary material, which is available to authorized users.

Introduction

Stigma originally was defined as “a dynamic process of devaluation that ‘significantly discredits’ an individual in the eyes of others” [1]. Since Goffman’s seminal work, research on stigma has explored negative social attitudes (perceived stigma), and discriminating behaviors (enacted stigma) toward stigmatized individuals [2,3]. Self-stigma (or internalized stigma) has also been examined, which results from fear of enacted stigma [4]. Some examples of self-stigma include individuals hiding their diagnoses or socially isolating themselves from others, or feeling ashamed of having a particular disease/condition.

The scientific literature on stigma has expanded considerably in the era of HIV/AIDS, given that having HIV disease is viewed more negatively than many other stigmatized conditions, such as mental illness and other physical health problems [5,6]. Conceptually, we know that HIV/AIDS stigma is multi-layered or compounded with already marginalized behaviors, such as sex work, drug use, and homosexual sexual practices; and vulnerable groups, such as prisoners or migrant populations [7]. We are also more aware of the many underlying factors at the community or social levels that may be causing HIV/AIDS stigma to surface, including lack of knowledge or understanding about the illness; misconceptions about how HIV is transmitted; lack of access to treatment; how media shapes the reporting on the epidemic; the incurability of AIDS; and existing prejudice and fear toward specific groups [7–9]. Furthermore, the concepts of perceived, enacted, internalized, and/or compounded stigma have been advanced by the development of quantitative measurements (some of them validated) to assess the types of stigma persons living with HIV/AIDS (PLWHA) are feeling or experiencing, or to assess non HIV-infected individuals’ perceived or enacted stigma towards PLWHA [10]. Lastly, a large body of research suggests that HIV/ AIDS stigma is a barrier associated with disclosure and negative health outcomes, impacting HIV preventive behaviors [8]; HIV care-seeking behaviors [11,12]; quality of care for PLWHA [11–13]; and treatment of PLWHA among loved ones [14] health care providers [12], and the larger community [8,9,15]. The negative health outcomes resulting from HIV/AIDS stigma have made it necessary for HIV prevention and treatment programs to focus greater efforts on reducing HIV/AIDS stigma.

Despite its negative impact on PLWHA, there surprisingly have been a limited number of intervention studies aimed at reducing HIV/AIDS stigma. What have been published, to date, are two reviews that summarize interventions with stigma reduction components [16,17], and one conceptual article describing what we can learn from HIV/AIDS stigma reduction interventions [9]. The 2003 review is the most relevant because of its focus on HIV/ AIDS stigma [16]. This review identified 22 intervention studies that had HIV/AIDS stigma reduction component(s), with the most common measurements of HIV/AIDS stigma being perceived stigma of non-infected populations towards PLWHA, and stigma perceived by PLWHA emanating from non-infected populations. In addition, this review provided a useful categorization of intervention strategies commonly implemented to reduce HIV/AIDS stigma. A second review was conducted on stigma reduction strategies for different diseases/ conditions [17]. Although it did not focus on HIV/ AIDS exclusively, this review was useful because it described similar stigma reduction strategies targeted at the individual, interpersonal, and community levels.

From these reviews, we have a better understanding of how HIV/AIDS stigma is being measured in intervention studies, types of stigma reduction strategies being implemented, and at which levels these stigma reduction strategies are being targeted. What is lacking from these reviews, however, is a *systematic* review to determine the quality of these studies vis-à-vis their effectiveness in reducing HIV/AIDS stigma. Preferred Reporting Items for Systematic reviews and Meta-Analyses (PRISMA) guidelines indicate that quality, or the

internal validity, should be assessed for each intervention study by examining the risk of bias at the study design and outcome levels [18]. In reporting this information for a given study, we can be more confident that the study's estimate of effect is correct. Our goal in the current systematic review is to update and expand upon the two prior reviews by broadening our search for intervention studies that may or may not target HIV/AIDS stigma, but measure HIV/AIDS stigma, pre-and post-intervention. This will potentially increase the pool of HIV-related interventions found since there are not as many that solely target stigma reduction. Another distinct feature of this review is that we will assess the risks of bias at the study design and outcome levels by rating the quality of the intervention studies reviewed. We will conclude by making recommendations on how to use this review to inform future HIV-related research/programs that incorporate HIV/AIDS stigma reduction as one of their objectives.

Methods

Search Strategy and Data Sources

To answer our research question, do populations that receive an HIV-related intervention have lower levels of HIV/AIDS stigma at posttest assessment, compared to pretest assessment?, our Population Intervention Comparison Outcome (PICO) search strategy [19] included a combination of key words and/or MESH terms to maximize our ability to find relevant articles (Supplemental Electronic Information 1).

With our last search occurring in March 2009, we identified published studies in the electronic databases of PubMed, PsychInfo, CINAHL, Social Work Abstracts, Web of Science (ISI), and the gray literature (unpublished research reports that are not found in traditional peer-reviewed publications) electronic databases of Google Scholar and Aegis. We also searched in NC Live, which is comprised of a number of databases, including gray literature databases.

We did not restrict publication date or the study population in which the intervention took place, but did limit the search to English-language articles and human populations. Moreover, unlike the 2003 review [16], we did not restrict the intervention to include a stigma reduction component in our search strategy, since our research question was interested in whether or not there was a reduction in HIV/AIDS stigma regardless of the HIV-related intervention used.

Study Selection

For the initial search, two reviewers independently reviewed the abstracts of studies to accept or reject for full text review using our PICO search strategy. Abstracts were rejected if the studies did not have (1) interventions that were HIV-related, (2) at least one *quantitative* HIV/AIDS stigma outcome measure, and (3) a study design that had pre-/post-intervention data collection time points. Any disagreements for which to accept or reject were resolved by discussion, and referral to a third reviewer, if necessary. For studies that were accepted at the abstract level, we did an ancestry search of their references to identify other relevant studies [20].

Full Text Review and Data Extraction

The same two reviewers independently reviewed the full texts of the studies identified from the electronic and ancestry searches to determine if they were still eligible to undergo data extraction. Data were extracted from eligible studies into an electronic spreadsheet. Reviewers met to reconcile any disagreements in the data extracted, and referral to third reviewer was done to resolve any disputes.

We extracted the following data: study characteristics (publication date, authors, study objectives, study population, study location, duration of intervention, description of the stigma measure, and the stigma outcome findings); type of study design; and intervention characteristics (types of intervention strategies used, whether or not part or all of the intervention focused on reducing HIV/AIDS stigma). To classify the intervention strategies, we used Brown et al.'s [16] scheme that was divided into four types of intervention strategies: (1) information-based approaches (e.g., written information in a brochure), (2) skill building (e.g., hands-on learning strategies for resolving negative attitudes), (3) counseling approaches (e.g., PLWHA support groups and resources), and (4) contact/interaction with PLWHA (e.g., testimonials, interaction between PLWHA and the general public).

In addition, we developed a checklist informed by a comparative effectiveness guide developed by the Agency for Healthcare Research and Quality (AHRQ) to assess the studies' internal validity and overall quality [21]. For the randomized control trials (RCTs) and pretest-posttest with non-randomized control group studies, we looked for any information that was given about (1) if and how participants were randomized; (2) the extent to which there was a risk for selection bias when comparing groups at baseline and loss to follow-up; (3) possible risk of confounding bias if the groups were not comparable at baseline, analysis was not intention-to-treat (ITT) (applicable for RCTs only), and/or confounders (typically would be certain demographic that were not similar among groups, but also could be other variables) were not adjusted for in the statistical analysis; and (4) possible risk of measurement bias if an un-validated measure of stigma was used, and the outcome assessor—the person or people who were responsible for analyzing the data—were not blinded to the assignment or exposure group status. For the pretest-posttest one group studies, we considered their internal validity inherently lowered because they do not have a control group. We further assessed these studies' internal validity by taking into consideration loss to follow-up, if confounders were identified and adjusted for in the statistical analysis, and whether or not a validated stigma measure was used. Using AHRQ's guidelines,[21] qualitative ratings of "good," "fair," or "poor" were given for each of these internal validity indicators to indicate low risk of bias (good) to high risk of bias (poor). We also rated for the extent to which the intervention focused on reducing HIV/AIDS stigma (All components = good, 1 or more components = fair, None = poor), and the statistics used and reported to demonstrate effectiveness (used statistics appropriately = good, did not use statistics appropriately = poor). We rated the studies on all of these variables, and then provided an overall quality rating of "good," "fair," or "poor," with "good" indicating the highest quality studies in this review (Supplemental Electronic Information 2).

Results

We selected 19 studies with HIV-related interventions that had an HIV/AIDS stigma outcome to answer our research question, do populations that receive an HIV-related intervention have lower levels of HIV/AIDS stigma at posttest assessment, compared to pretest assessment? Following PRISMA guidelines [18], Fig. 1 illustrates our process for selecting studies that were included in this systematic review.

Our initial search identified 515 abstracts. After reviewing these abstracts, only 35 studies were accepted, based on our PICO search strategy. We performed an ancestry search of these 35 studies' references to identify any other relevant studies. The ancestry search yielded 40 additional studies that we selected by their titles. We then full text reviewed the 35 studies from the electronic search, and the 40 studies from the ancestry search, to determine if they were still eligible. Only 22 of these studies were still eligible after full text review, and were selected for data extraction. The reason for rejecting the other 53 studies

primarily was because they did not meet our study design criterion after reading their full texts. Furthermore, during our data extraction, we discovered and rejected 3 of the 22 studies that did not use paired data to appropriately evaluate pre- and post-intervention changes in their HIV/AIDS stigma outcome. Thus, 19 studies were included in this systematic review, seven of which were not part of Brown et al.'s [16] review given that they were published in 2004 or later. While the majority of the studies were published and found in one of the electronic databases, one unpublished dissertation [22] was found in Proquest (one of the databases covered in NC Live) and two published studies [23,24] were found through our ancestry search.

Study and Intervention Characteristics

Table 1 (and Supplemental Electronic Information 3) summarizes the study and intervention characteristics by study design for the 19 studies. [22–40] Nine of the studies were RCTs [25–33], six were non-randomized control group study designs [24,34–38], and four were pretest/ posttest without a control group study designs [22,23,39,40]. Of the 9 RCTs, six were conducted in the United States or in other North American or European countries [26,27,29–32], two in Asian countries [25,33], and one in an African country [28]. Four of the RCTs focused on student populations [27–29,31], two focused on health care provider populations [32,33], one was a community level intervention [25], and the remaining two focused on families [30] and women [26]. Their interventions' durations ranged from 15 minutes to 2 years. We then used Brown et al.'s classification of types of interventions to categorize the intervention components [16]. All nine of the studies used informational approaches to deliver their intervention. One of the studies used all four of the intervention strategy categories (informational approaches, skill building, counseling/support, and PLWHA testimonials) [25]. Two of the studies used 3 of the 4 strategies (informational approaches, skill building, and PLWHA testimonials) [31,33].

Of the six non-randomized control group studies, three were conducted in African countries [34,35,38], two were conducted in Asian countries [24,36], and one was conducted in a European country [37]. Three of these studies focused on student populations [34,35,37], one was conducted with health care providers [36], one was a community level intervention [24], and the last focused on working women [38]. Their interventions' duration ranged from 3 days to 1 year. Like the RCTs, all six of the non-randomized control group studies used informational approaches to deliver their intervention. One of the studies used three of the strategies (informational approaches, skill building, and PLWHA testimonials) [36], and two of the studies additionally used skill building [34] and support groups [24] along with informational approaches to deliver their interventions.

For the four studies without a control group, three were conducted in the United States [22,23,40], and one was conducted in India [39]. Three of the studies focused on health care providers [22,39,40], and the fourth study focused on PLWHA [23]. Their interventions' duration ranged from 1 hour to 3 years. All four of these studies used informational approaches to deliver their intervention. Two of the studies additionally used counseling/ support groups to deliver their interventions [23,39]. One of the studies used all four of the intervention strategies [39].

Description of Stigma Outcome Measures and Findings

Table 2 (and Supplemental Electronic Information 4) organizes the 19 studies by the extent to which their interventions focus on stigma reduction, and for each study, presents the study population(s) targeted, a brief description of the types of HIV/AIDS stigma measures used (number of items, whether or not they were validated, and types of stigma measured), and

the findings associated with the measures. Interestingly, none of the 19 studies used a common measure to evaluate pre- and post-changes in HIV/AIDS stigma.

All 19 studies measured perceived stigma, mainly in the form of asking respondents to agree or disagree on statements reflecting various attitudes toward PLWHA. Seven of these 19 studies also measured enacted stigma, mainly in the form of asking respondents whether they agree or disagree with different discriminatory behaviors towards PLWHA [24,25,30,33,35,36,39]. None of the 19 studies measured internalized or compounded stigma.

Nine of the 19 studies used either a previously/originally developed and validated, multi-item scale/index in which a summary or mean score was calculated, ranging from 4 to 30 items [22,23,25,27–30,32,38]. Of these nine studies, however, three did not provide a score range and its interpretation for their measures to adequately assess if HIV/AIDS stigma was reduced or had increased as a result of the intervention [22,27,32]. Two of the studies used scales to measure stigma from the perspectives of PLWHA [23,25], and only one of these studies measured stigma in both un-infected and PLWHA participants [25]. Of the nine studies that used validated scales, seven studies reported a statistically significant post-intervention change in their HIV/AIDS stigma outcome measure at $P < .05$ or lower [23,25,27–30,38], but only 5 of these 7 studies convincingly demonstrated post-intervention stigma reduction, based on their reported stigma outcome results [23,25,27,28,30]. The other two studies that were statistically significant did not report the actual results (e.g., effect size or score differences) [29], or obtain an effect size >1 for either the intervention or control group [38].

Ten of the 19 studies measured HIV/AIDS stigma with originally developed and un-validated individual items, not constituting a scale, with results reported for each individual item. Six of the 10 studies measured HIV/AIDS stigma with 1–3 items [26,28,31,33,35,37], and four studies with greater than four items [24,36,39,40]. For one of these studies with 12 items, score differences were not reported for its control group [24]. Only one of these studies measured stigma in both un-infected and PLWHA participants [24]. Of the 10 studies that used individual items to measure HIV/AIDS stigma, six studies reported a statistically significant post-intervention change in one or more of their HIV/AIDS stigma items at $P < .05$ or lower [24,34–37,39]. For two of the studies, only 1 of 3 of their items demonstrated reduced HIV/AIDS stigma post-intervention [35,37]. For the study that measured 12 items, 11 of the 12 items demonstrated stigma reduction in the intervention group, but these items' pre/post-intervention differences were not reported for the control group, making it difficult to evaluate this study based on its original study design (non-randomized control group) [24].

Quality Assessment

In Table 3, we present the 19 studies' quality assessment using the checklist in Supplemental Electronic Information 2. The 19 studies were categorized by study design, and whether or not they had a statistically significant reduction in their HIV/AIDS stigma outcome. None of the 19 studies reported power calculations to determine if they had sufficient samples sizes to assess effectiveness of their interventions.

RCTs' Quality Ratings

Of the nine RCTs, five had a statistically significant reduction in their stigma outcome measurement [25,27–30] (Table 3). Only 2 [25,33] of the 9 RCTs had interventions that focused solely on reducing HIV/AIDS stigma. The majority of the nine RCTs did not identify their method of randomization, nor if they concealed the order of assignments from

investigators involved in these trials, and thus, were given “poor” ratings for these factors. Ashworth [26]—a non-significant study—however, did mention a method of randomization, and was given a “fair” rating on this factor.

Degree of selection bias relating to comparison of groups at baseline and loss to follow-up varied in the RCTs. For the significant studies, Krauss [30] minimized selection bias by demonstrating both similar baseline characteristics in groups and low differential loss to follow-up. Apinundecha [25] also demonstrated similar baseline characteristics in groups, but failed to report any information to evaluate if there was loss to follow-up, and thus, was given a “poor” rating for this factor. For the non-significant studies, Wu [33] had “good” ratings for selection bias, followed by Ashworth [26] that was given a “fair” rating for similar baseline characteristics in groups since this study mentioned having similarities without presenting a table to demonstrate these similarities.

Degree of confounding bias was rated based on if they stated that they used an Intention to Treat (ITT) analysis (takes into account initial assignment and any loss to follow-up), and any adjustment of confounders if there were differences in baseline characteristics (usually demographics, but could be other variables) in groups. For the significant studies, only Krauss [30] specified doing both an ITT and cross-over analysis, and thus, received “good” ratings for confounding bias. Apinundecha [25], Klepp [28], and Krauss [30] received “good” ratings for controlling for confounders; Apinundecha [25] and Krauss [30] through their study design, and Klepp [28] through controlling for some confounders in their statistical analysis.

Degree of measurement bias was rated based on whether or not a validated HIV/AIDS stigma measurement—preferably a scale that measured different stigma constructs—was used, and if the outcome assessors were blinded to the group assignment. None of the RCTs reported having a blinded outcome assessor, and were given “poor” ratings for this indicator. For the significant RCTs, only Apinundecha [25] and Krauss [30] used validated instruments to measure perceived and enacted HIV/AIDS stigma, but Apinundecha’s [25] instrument measured stigma in both HIV-infected and un-infected members of their study population.

Lastly, 6 of the 9 RCTs [25,26,30,31,33] demonstrated appropriate statistical testing and reported either effect sizes, or mean score/% differences with associated p-values (none of them reported confidence intervals). Based on the internal validity factors, the interventions’ focus, and the statistics used and reported, Krauss [30], Apinundecha [25] and Wu [33] received “good” overall quality ratings. Wu’s [33] non-significant stigma finding may have been due to insufficient sample size or inadequate measures of stigma, rather than something related to how they conducted their intervention trial.

Non-Randomized Studies’ Quality Ratings

Table 3 also presents the quality assessments for the non-randomized control group studies and pretest–posttest one group studies. All of the six non-randomized control group studies demonstrated significant results for their HIV/AIDS stigma outcome [24,34–38]. Only Yang [24] had an intervention that solely focused on reducing HIV/AIDS stigma. For the four pretest–posttest one group studies, only three [23,39,40] demonstrated significant results for their HIV/AIDS stigma outcome, and two of these studies [27,28] had interventions with one or more components focused on reducing HIV/AIDS stigma.

For the non-randomized control group studies, we gave them “poor” ratings for the randomization indicators because they did not use randomization to assign participants to

intervention and control groups. For the pretest–posttest one group studies, the randomization indicators were not applicable.

Degree of selection bias for the non-randomized control group studies was based on the same criteria as the RCTs, but only loss to follow-up was considered for the pretest–posttest one group studies. For the non-randomized control group studies, only Fawole [34] had “good” ratings for selection bias. For the pretest–posttest one group studies, only Zachariah [22] received a “good” rating by demonstrating no loss to follow-up.

Degree of confounding bias only was rated based on if they identified potential confounders (e.g., demographics), and controlled for these confounders in their statistical analyses. Only Fawole’s [34] study adequately addressed confounders in their statistical analysis and was given a “good” rating.

Degree of measurement bias for the non-randomized control group studies was based on the same criteria as the RCTs, but for the pretest–posttest one group studies, only the use of an un-validated HIV/AIDS stigma measurement was considered. For the non-randomized control group studies, only Fawole [34] reported having the outcome assessors blinded to exposure status of group, and was given a “good” rating for this indicator. All of the non-randomized control group studies, however, were given “fair” or “poor” ratings related to the stigma outcome measurement they used. Fawole [34], Kuhn [35], Lueveswanij [36], Merakou [37], and Yang [24] used un-validated, originally developed instruments that measured primarily attitudes toward PLWHA of un-infected study populations; in the case of Fawole [34], Kuhn [35], Lueveswanij [36], and Merakou [37], between 1 and 4 items constituted their stigma measurement. For the same reasons, we gave “fair” or “poor” ratings for the stigma outcome measurements used to all of the pretest–posttest one group studies.

Lastly, only Fawole [34] from the non-randomized control group studies, and Pital [39] from the pretest–posttest one group studies were given “good” ratings for demonstrating appropriate statistical testing and reporting either effect sizes, or mean score/% differences with associated p-values (none of them reported confidence intervals). Based on the internal validity factors, the interventions’ focus, and statistics used and reported, only Fawole’s [34] study had a “fair” overall quality rating. All of the pretest–posttest one group studies were considered “poor.”

Discussion

This systematic review revisits the question on the availability of effective interventions to reduce HIV/AIDS stigma given that HIV/AIDS stigma results in negative health outcomes in both industrialized and developing contexts. Of the 19 studies reviewed, 14 demonstrated effectiveness in reducing HIV/AIDS stigma [23–25,27–30,34–40]. We also assessed the quality of these 19 studies by rating their internal validity, the extent to which their interventions targeted stigma reduction, the HIV/AIDS stigma measures used, and the statistics reported to demonstrate pre- and post-intervention changes in their HIV/AIDS stigma outcomes. We then applied an overall quality rating to each study. Only 2 of the 14 effective studies—Apinundecha [25] and Krauss [30]—were given overall “good” quality ratings. Although not significant, Wu’s [33] RCT provider intervention in China also received an overall “good” rating. The paucity of good quality studies within the last 20 years identified in this review reveals the current gaps in evidenced-based interventions to reduce HIV/AIDS stigma. These gaps include (1) not enough interventions targeting HIV/AIDS stigma, (2) using disparate and inadequate measures to evaluate HIV/AIDS stigma

reduction, (3) whether statistically significant stigma reduction in these intervention trials has public health significance, and (4) the lack of good quality (internally valid) studies.

Regarding the first gap, only three of the studies in this review—Apinundecha [25], Yang [24], and Wu [33]—tested interventions that aimed to reduce HIV/AIDS stigma. This was surprising given the vast literature on HIV/AIDS stigma's negative impact on HIV prevention and treatment efforts, and repeated recommendations by stigma experts to target HIV/AIDS stigma in their interventions and programs [12,41]. Furthermore, Apinundecha's [25] and Wu's [33] studies had all, or the majority of the intervention strategies (information, skill building, counseling, and PLWHA testimonials) identified from Brown et al.'s [16] review incorporated into their interventions. The combination of developing interventions that focus on stigma reduction, and using more than one strategy to deliver the interventions, may have greater impact to enhance participants' understanding about the effect of HIV/AIDS stigma in HIV-infected and/or at-risk populations.

This review brings attention to a second gap relating to the disparate and/or inadequate measures used to evaluate stigma reduction in HIV intervention trials. The measures used are disparate because none of the 19 studies in this review used a common scale despite the fact that a recently published systematic review identified 23 HIV/AIDS stigma instruments that are evidenced-based and available to researchers [10]. Only the scale used by Apinundecha [25] was identified in the aforementioned systematic review, and has been widely used in other studies [10]. The measures used in the 19 studies reviewed also are inadequate, based on their lack of internal and/or external validation, and their narrow focus on assessing, primarily, perceived stigma in un-infected populations. Only nine of the studies in this review used a previously or originally validated scale to measure HIV/AIDS stigma [22,23,25,27–30,32,38], and of these studies, only Apinundecha [25] used a measure that is a multi-item, externally validated scale that measures community-level perceived and enacted stigma [42].

Along with measurement issues is a third gap associated with whether the intervention studies that demonstrated statistical significance in reducing HIV/AIDS stigma have public health significance. For public health significance, we would like to see whether a reduction in HIV/AIDS stigma was associated with improved change in health outcomes, such as increased HIV testing, increased access to or utilization of HIV treatment/care, improved policy support for PLWHA, improved mental health outcomes and social support, and/or reduced HIV-related symptoms [10]. None of the 14 statistically significant studies examined the relationship between reducing HIV/AIDS stigma and any health outcomes associated with HIV prevention and treatment. Krauss's RCT [30], however, may have demonstrated some public health relevance since the improved change in comfort level to interact with PLWHA was twice as high in the intervention group versus the control or crossover groups. For the statistically significant studies that did not use a validated scale, and/or used one to a few items to measure one or more of the complex constructs of HIV/AIDS stigma [34–37], the evidence of stigma reduction is generally less meaningful.

The last gap highlighted in this review relates to the methodological rigor of the intervention trials themselves. Typically, sample size calculations would be performed and reported for any study in which effects of a particular outcome are being measured. None of the 19 studies reported sample size calculations, and for the non-significant studies, the negative findings could be a result of inadequate sample sizes. This is particularly the case with Wu's study, which seemed to have fairly good internal validity, an intervention that targeted stigma reduction, and appropriate statistical testing, but its pre/post-intervention findings were not significant [33]. It is unclear if the negative findings were a result of the poor stigma measure that they used, or because of inadequate sample sizes. Second, while using

an RCT study design—when conducted properly—has an inherent advantage to minimize selection and confounding biases that can affect intervention trials, all nine of the RCT studies reviewed did not use “good” methods of randomization and allocation concealment when assigning individuals or communities to groups. A consequence of poor randomization is dissimilar groups at baseline that 6 of the 9 RCTs exhibited, and were given fair to poor ratings for this quality indicator [26–29,31,32]. Third, the most appropriate analytic approach in RCTs would be intention-to-treat (ITT) to take into account participants who were lost to follow-up to further minimize confounding bias. Only Krauss reported both ITT and crossover analyses to account for parents who declined to participate in the intervention [30]. Lastly, having an outcome assessor (e.g., statistician) who is blinded to the group—in either RCTs or non-randomized control group designs—is helpful to reduce measurement bias. None of the RCTs reported having an outcome assessor blinded to group assignment, and only Fawole [34] of the six non-randomized group studies reported using a blinded outcome assessor. Using other study designs besides an RCT to assess effectiveness in reducing stigma is much more challenging, and this review demonstrates the overall poorer quality of the other 10 studies that used non-randomized control group or pretest/posttest one group study designs.

Like other qualitative reviews of intervention studies, this systematic review has its limitations. The gold standard for reviewing intervention effectiveness is with a meta-analysis that calculates a pooled effect size from RCTs. In this review, the stigma instruments of the RCTs were too heterogeneous to combine in a meta-analysis. Thus, we broadened our criteria to examine three different types of study designs typically used to conduct public health/social science intervention trials, and used a checklist to rate the overall quality of each of the studies adapted from *AHRQ's Methods Guide for Comparative Effectiveness* [21] that satisfies PRISMA guidelines for conducting systematic reviews [18].

To address the gaps identified in this review, we have several recommendations that may help guide researchers and organizations to develop studies in the future that better determine whether interventions are effective or not in reducing HIV/AIDS stigma. First, interventions that are not focused on reducing stigma, but that demonstrate a reduction in HIV/AIDS stigma assessed with one or a few items, is inadequate. Designing stigma reduction interventions should include formative research to understand the issues of HIV/AIDS stigma affecting a particular target population, and using the formative data to develop intervention components that target the issues of HIV/AIDS stigma identified. Second, when considering which instrument to use to measure changes in HIV/AIDS stigma, it is important to choose an instrument that has undergone internal validation (e.g., psychometric analysis), and possibly external validation of its items; these instruments will more likely be measuring more than one concept (i.e., perceived, enacted, internalized, or compounded stigma). Third, researchers should consider using RCT study designs, if possible, to conduct their intervention trials, paying attention to issues pertaining to internal validity. Lastly, even with a well-designed intervention that demonstrates a reduction in HIV/AIDS stigma, we currently know very little about how or if the change in stigma reduction affects associated health outcomes, such as increasing HIV testing, increasing access to HIV treatment/ care, improving policy support for PLWHA, improving mental health outcomes and social support, and/or reducing HIV-related symptoms [10]. Thus, future research on the effectiveness of HIV/AIDS stigma reduction interventions must additionally measure health outcomes of interest, pre-/post-intervention, to understand and appreciate the need for HIV/AIDS stigma reduction globally.

Conclusion

In conclusion, updating our current knowledge about interventions to reduce HIV/AIDS stigma is an important step in understanding the breadth of HIV/AIDS stigma reduction efforts currently available. Of the 19 studies reviewed, only three RCTs (two significant and one not significant) were considered of higher quality, and that we would recommend to be replicated in similar and other cultural contexts [25,30,33]. More importantly, we hope that our recommendations to guide future intervention studies targeting HIV/AIDS stigma reduction will be useful to researchers and organizations dedicated to this work.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

This study was funded by the National Institute of Nursing Research, Grant #R01NR010204-01A2; and National Institute of Allergy and Infectious Diseases, Grant #K01AI055247-05. This project also was supported by Award #UL1RR025747 from the National Center for Research Resources, with the content being the sole responsibility of the authors and not necessarily representing the official views of the National Center for Research Resources or the National Institutes of Health. We would like to acknowledge Michelle Hayes, Eburn Odeneye, and Nana Abrew of the UNC Cecil G. Sheps Center for Health Services Research for assisting us with compiling the articles for data extraction. We also would like to thank Dr. Karen Crotty for her review of drafts of the manuscript.

References

1. Stigma, GE. Notes on the management of spoiled identity. Englewood Cliffs, NJ: Prentice Hall; 1963.
2. Corrigan PW, Penn DL. Lessons from social psychology on discrediting psychiatric stigma. *Am Psychol.* 1999; 54(9):765–76. [PubMed: 10510666]
3. Pryor JB, Reeder GD, Yeadon C, Hesson-McLennis M. A dual-process model of reactions to perceived stigma. *J Pers Soc Psychol.* 2004; 87(4):436–52. [PubMed: 15491270]
4. Rüsch N, Angermeyer MC, Corrigan PW. Mental illness stigma: concepts, consequences, and initiatives to reduce stigma. *Eur Psychiatry.* 2005; 20(8):529–39. [PubMed: 16171984]
5. Corrigan PW, River LP, Lundin R, Wasowski K, Campion J, Mathisen J, et al. Stigmatizing attributions about mental illness. *J Community Psychol.* 2000; 28(1):91–102.
6. Crawford AM. Stigma associated with AIDS: a meta-analysis. *J Appl Soc Psychol.* 1996; 26(5): 398–416.
7. HIV-related stigma, discrimination and human rights violations: Case studies of successful programmes. Joint United Nations Programme on HIV/AIDS (UNAIDS); 2005 [Accessed 28 Oct 2010]. Available at: http://data.unaids.org/publications/irc-pub06/jc999-humrightsviol_en.pdf
8. Sengupta S, Strauss RP, Miles MS, Roman-Isler M, Banks B, Corbie-Smith G. A conceptual model exploring the relationship between HIV stigma and implementing HIV clinical trials in rural communities of North Carolina. *N C Med J.* 2010; 71(2):113–22. [PubMed: 20552760]
9. Mahajan AP, Sayles JN, Patel VA, et al. Stigma in the HIV/AIDS epidemic: a review of the literature and recommendations for the way forward. *AIDS.* 2008; 22(Suppl 2):S67–79. [PubMed: 18641472]
10. Earnshaw VA, Chaudoir SR. From conceptualizing to measuring HIV stigma: a review of HIV stigma mechanism measures. *AIDS Behav.* 2009; 13(6):1160–77. [PubMed: 19636699]
11. Prachakul W, Grant JS, Keltner NL. Relationships among functional social support, HIV-related stigma, social problem solving, and depressive symptoms in people living with HIV: a pilot study. *J Assoc Nurses AIDS Care.* 2007; 18(6):67–76. [PubMed: 17991600]
12. Sayles JN, Ryan GW, Silver JS, Sarkisian CA, Cunningham WE. Experiences of social stigma and implications for healthcare among a diverse population of HIV positive adults. *J Urban Health.* 2007; 84(6):814–28. [PubMed: 17786561]

13. Vanable PA, Carey MP, Blair DC, Littlewood RA. Impact of HIV-related stigma on health behaviors and psychological adjustment among HIV-positive men and women. *AIDS Behav.* 2006; 10(5):473–82. [PubMed: 16604295]
14. Bogart LM, Cowgill BO, Kennedy D, et al. HIV-related stigma among people with HIV and their families: a qualitative analysis. *AIDS Behav.* 2008; 12(2):244–54. [PubMed: 17458691]
15. Foster PH. Use of stigma, fear, and denial in development of a framework for prevention of HIV/AIDS in rural African American communities. *Fam Community Health.* 2007; 30(4):318–27. [PubMed: 17873638]
16. Brown L, Macintyre K, Trujillo L. Interventions to reduce HIV/AIDS stigma: what have we learned? *AIDS Educ Prev.* 2003; 15(1):49–69. [PubMed: 12627743]
17. Heijnders M, Van Der Meij S. The fight against stigma: an overview of stigma-reduction strategies and interventions. *Psychol Health Med.* 2006; 11(3):353–63. [PubMed: 17130071]
18. Moher D, Liberati A, Tetzlaff J, Altman DG, the PRISMA Group. Reprint—preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. *Phys Ther.* 2009; 89(9): 873–80. [PubMed: 19723669]
19. da Costa Santos CM, de Mattos Pimenta CA, Nobre MR. The PICO strategy for the research question construction and evidence search. *Rev Lat Am Enfermagem.* 2007; 15(3):508–11. [PubMed: 17653438]
20. Conn VS, Isaramalai SA, Rath S, Jantarakupt P, Wadhawan R, Dash Y. Beyond MEDLINE for literature searches. *J Nurs Scholarsh.* 2003; 35(2):177–82. [PubMed: 12854300]
21. Agency for Healthcare Research and Quality. *Methods Guide for Comparative Effectiveness Reviews, Version 1.0.* [Accessed 28 Oct 2010]. [Draft posted Oct. 2007]. Available at: http://www.effectivehealthcare.ahrq.gov/ehc/products/60/318/2009_0805_grading.pdf
22. Zachariah, GS. Dissertation presented to Spalding University. Louisville, KY: 1998. AIDS fear in health care workers: Development of educational program to decrease the fear.
23. Rounds KA, Galinsky MJ, Despard MR. Evaluation of Telephone Support Groups for persons with HIV disease. *Res Soc Work Pract.* 1995; 5:442.
24. Yang Y, Zhang K. HIV/AIDS-related discrimination in Shanxi rural areas of China. *Biomed Environ Sci.* 2004; 17:410–7. [PubMed: 15745245]
25. Apinundecha C, Laohasiriwong W, Cameron MP, Lim S. A community participation intervention to reduce HIV/AIDS stigma, Nakhon Ratchasima province, northeast Thailand. *AIDS Care.* 2007; 19(9):1157–65. [PubMed: 18058400]
26. Ashworth CS, DuRant RH, Gaillard G, Rountree J. An experimental evaluation of an AIDS educational intervention for WIC mothers. *AIDS Educ Prev.* 1994; 6(2):154–62. [PubMed: 7517156]
27. Gill NJ, Beazley RP. Grade 6 students benefit from learning about AIDS. *Can J Public Health.* 1993; 84(Suppl 1):S24–7. [PubMed: 8481863]
28. Klepp KI, Ndeki SS, Leshabari MT, Hannan PJ, Lyimo BA. AIDS education in Tanzania: promoting risk reduction among primary school children. *Am J Public Health.* 1997; 87(12):1931–6. [PubMed: 9431279]
29. Knaus CS, Austin EW. The AIDS Memorial Quilt as preventative education: a developmental analysis of the Quilt. *AIDS Educ Prev.* 1999; 11(6):525–40. [PubMed: 10693648]
30. Krauss BJ, Godfrey CC, O'Day J, Freidin E. Hugging my uncle: the impact of a parent training on children's comfort interacting with persons living with HIV. *J Pediatr Psychol.* 2006; 31(9):891–904. [PubMed: 16452647]
31. Markham C, Baumler E, Richesson R, et al. Impact of HIV-positive speakers in a multicomponent, school-based HIV/STD prevention program for inner-city adolescents. *AIDS Educ Prev.* 2000; 12(5):442–54. [PubMed: 11063063]
32. Stewart KE, DiClemente RJ, Ross D. Adolescents and HIV: theory-based approaches to education of nurses. *J Adv Nurs.* 1999; 30(3):687–96. [PubMed: 10499226]
33. Wu S, Li L, Wu Z, et al. A brief HIV stigma reduction intervention for service providers in China. *AIDS Patient Care STDS.* 2008; 22(6):513–20. [PubMed: 18462076]

34. Fawole IO, Asuzu MC, Oduntan SO, Brieger WR. A school-based AIDS education programme for secondary school students in Nigeria: a review of effectiveness. *Health Educ Res.* 1999; 14(5): 675–83. [PubMed: 10510075]
35. Kuhn L, Steinberg M, Mathews C. Participation of the school community in AIDS education: an evaluation of a high school programme in South Africa. *AIDS Care.* 1994; 6(2):161–71. [PubMed: 8061076]
36. Lueveswanij S, Nittayananta W, Robison VA. Changing knowledge, attitudes, and practices of Thai oral health personnel with regard to AIDS: an evaluation of an educational intervention. *Community Dent Health.* 2000; 17(3):165–71. [PubMed: 11108404]
37. Merakou K, Kourea-Kremastinou J. Peer education in HIV prevention: an evaluation in schools. *Eur J Public Health.* 2006; 16(2):128–32. [PubMed: 16431871]
38. Norr KF, Norr JL, McElmurry BJ, Tlou S, Moeti MR. Impact of peer group education on HIV prevention among women in Botswana. *Health Care Women Int.* 2004; 25(3):210–26. [PubMed: 15195767]
39. Pisal H, Sutar S, Sastry J, et al. Nurses' health education program in India increases HIV knowledge and reduces fear. *J Assoc Nurses AIDS Care.* 2007; 18(6):32–43. [PubMed: 17991597]
40. Sowell R, Seals G, Wilson B, Robinson C. Evaluation of an HIV/AIDS continuing education program. *J Contin Educ Nurs.* 1998; 29(2):85–93. [PubMed: 9582777]
41. Nyblade LC. Measuring HIV stigma: existing knowledge and gaps. *Psychol Health Med.* 2006; 11(3):335–45. [PubMed: 17130069]
42. Herek GM, Capitanio JP, Widaman KF. HIV related stigma and knowledge in the United States: prevalence and trends, 1991–1999. *Am J Public Health.* 2002; 92:371–7. [PubMed: 11867313]

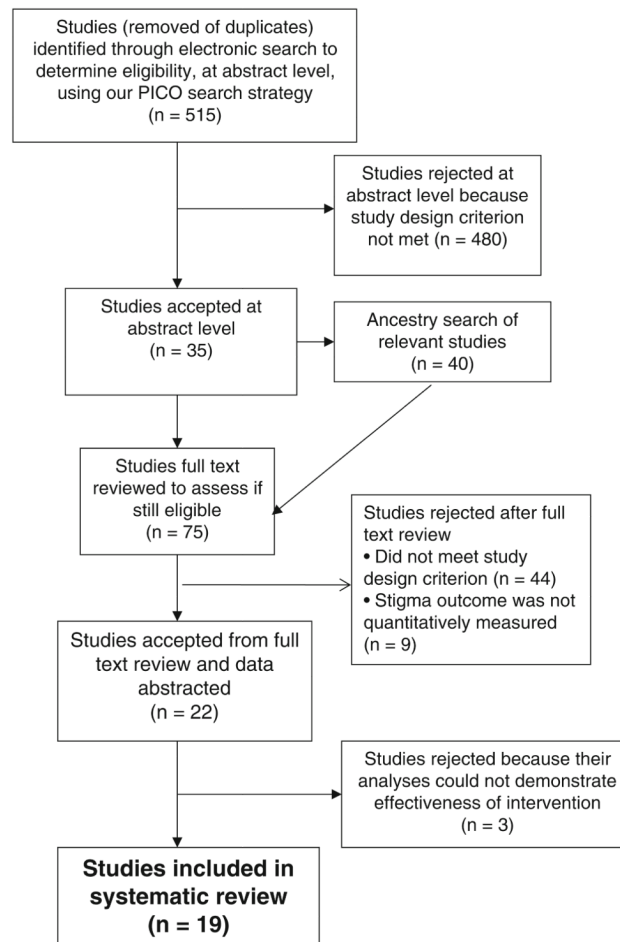


Fig. 1.
Selection process for studies included in systematic review

Table 1

Study and intervention characteristics of the 19 studies

1st Author, publication date	Study population	Baseline sample size	Intervention strategies^a
Country		# Arms	Intervention duration
RCTs			
Apinundecha [25], 2008	Community	66	I, SB, CS, T
Thailand		2	8 months
Ashworth [26], 1994	Women in WIC program	217	I
United States		3	15–18 min
Gill [27], 1993	Students (6th grade)	257	I
Canada		2	200 min
Klepp [28], 1997	Students (6th grade)	1,063	I
Tanzania		2	20 h
Knaus [29], 1999	Students (College)	416	I
United States		2	1-time attending quilt display
Krauss [30], 2006	Parents and pre-adolescent children	238	I
United States		3	12 h
Markham [31], 2000	Students (High school)	1,491	I, SB, T
United States		2	2 years
Stewart [32], 1999	Providers (Nurses)	88	I
United Kingdom		2	90 min
Wu [33], 2008	Providers	138	I, SB, T
China		2	4 h
Non-randomized control group studies			
Fawole [34], 1999	Students (High school)	450	I, SB
Nigeria		2	6 weeks
Kuhn [35], 1994	Students (High school)	567	I
South Africa		2	2 weeks
Lueveswanij [36], 2000	Providers (Oral health)	139	I, SB, T
Thailand		2	3 days
Merakou [37], 2006	Students (High school)	702	I
Greece		2	1 year
Norr [38], 2004	Working women	403	I
Botswana		2	6, 90-min sessions
Yang [24], 2004	Community	490 ^b	I, CS
China		1 ^b	1 year
Pretest/posttest one group studies			
Pisal [39], 2007	Providers (Nurses)	371	I, SB, CS, T
India		1	4 days
Rounds [23], 1995	PLWHA	18	I, CS
United States		1	6 h, 1 h/week
Sowell [40], 1998	Providers (Nurses)	173	I

1st Author, publication date	Study population	Baseline sample size	Intervention strategies^a
Country		# Arms	Intervention duration
United States		1	3 years
Zachariah [22], 1998	Providers (Nurses)	35	I
United States		1	1 h

^aBrown et al.'s [25] classification of types of intervention strategies: informational approaches (I), skill building (SB), counseling/support (CS), and PLWHA testimonials (T)

^bFor the Yang [24] study, a non-randomized control group study design was reported, but the baseline sample size (and subsequent findings) only was reported for the intervention group

Table 2

Stigma measures' description and findings of the 19 studies

1st Author study design ^a	Administered to: HIV-, HIV?, both	Validated/un-validated # Items Types of stigma ^b	Results or NS ^c	Interpretation
Entire intervention targeting HIV/AIDS stigma reduction				
Apinundecha [25] RCT	Both	Validated 30-item scale PS, ES	ANCOVA score differences, $P < .01$ Intervention: 35.1 Control: .98	Intervention had reduced stigma; higher score represents lower stigma, post-intervention
Yang [24] Nonrandom grp.	Both	Un-validated 12 items PS, ES	Chi-square, 11/12 items, $P < .01$ 5–20% increase in “yes” responses after intervention; no data reported for control group	Increase in “yes” responses represented less stigma, post-intervention
Wu [33] RCT	HIV–	Un-validated 3 items PS, ES	Mixed effects models: NS	No group differences on agreement/ item = no differences in stigma
1 or more components of intervention targeting HIV/AIDS stigma reduction				
Gill [27] RCT	HIV–	Validated 14-item scale PS	Paired t -tests, mean %, $P < .001$ Intervention: 81% Control: 71%	Intervention group had more positive attitudes toward PLWHA, post- intervention
Klepp [28] RCT	HIV–	Validated 4-item scale PS	Mixed Model ANOVA, $P = .0015$ Effect size = 2.8	Intervention group had more positive attitudes toward PLWHA, post- intervention
Krauss [30] RCT	HIV–	Validated 22-item sub-scale PS, ES	ANCOVA, score difference, $P = .001$ Intervention: 6.0 Control: 3.0 Crossover: 3.0	Children's total comfort in interacting with PLWHA increased most in intervention group, post-intervention
Lueveswanij [36] Nonrandom grp.	HIV–	Un-validated 4 items PS, ES	Paired t -test/item, differences, $P < .05$ Intervention: 5–50% Control: –5–3%	Intervention group had more positive attitudes toward PLWHA, post- intervention
Markham [31] RCT	HIV–	Un-validated 2 items PS	Generalized linear regression model: NS for effect of overall intervention	Unsure what NS means since no response options reported for items
Norr [38] Nonrandom grp.	HIV–	Validated (4 of 6 items) 6-item scale PS	Paired t -tests, mean scores, $P < .001$ Intervention: 5.0 Control: 4.1	Intervention group had more positive attitudes toward PLWHA, but < 1 effect size at follow-up
Pisal [39] One group	HIV–	Un-validated 11 items PS, ES	Paired t -tests/item, $P < .0001$: score differences range from 9.4–57.5	10/11 items had decreased scores, post- intervention, representing less stigma
Sowell [40] One group	HIV–	Un-validated 7 items PS	Reported pre/post item results as significant, but no test statistic/ p - values: improvement ranged from 2 to 23%	Pre/post improvement in attitudes toward caring for PLWHA
No intervention components targeting HIV/AIDS stigma reduction				
Ashworth [26] RCT	HIV–	Un-validated 1 item PS	Kruskall Wallis ANOVA: NS	No effect of intervention on stigma item
Fawole [34] Nonrandom grp.	HIV–	Un-validated 1 item PS	ANOVA, % agree on item, $P < .05$ Intervention: 79% ^d Control: 14% ^d	Intervention group had more % agreement on if they could touch and care for PLWHA at post-intervention

1st Author study design ^a	Administered to: HIV-, HIV?, both	Validated/un-validated # Items Types of stigma ^b	Results or NS ^c	Interpretation
Knaus [29] RCT	HIV-	Validated 8-item index (6 items on stigma) PS	Paired <i>t</i> -tests, mean score, <i>P</i> = .025 Intervention: 35.3 Control: 34.3	Intervention group had more positive attitudes toward PLWHA, but score difference was small, post-intervention
Kuhn [35] Nonrandom grp.	HIV-	Un-validated 3 items PS, ES	Test statistic not reported, 1 item, <i>P</i> = .0001 Intervention: 41.2% Control: 10.8%	Intervention group would accept PLWHA student into their class more than control group at post-intervention
Merakou [37] Nonrandom grp.	HIV-	Un-validated 3 items PS	Chi-square test/item—1 item significant Intervention: 9%, <i>P</i> = .006 Control: 16%, .0005	Intervention group had lower agreement that HIV carriers should not have sexual contacts at post-intervention
Rounds [23] One group	HIV+	Validated 8-item scale PS	Wilcoxon matched pairs signed-ranks test, pre/post score difference <i>P</i> < .05: 1.78	Decreased feelings of being isolated or left out by others, post-intervention
Stewart [32] RCT	HIV-	Validated 10-item sub-scale PS	2 × 3 MANCOVA: NS	No intervention effect on stigma
Zachariah [22] One group	HIV-	Validated 2 scales on AIDS phobia and homophobia (# items not reported) PS	Paired <i>t</i> -tests reported/scale: NS	No intervention effect on AIDS phobia or homophobia

^aFor study designs, “Nonrandom grp.” is abbreviated for pre-test/post-test with non-randomized control group; “One group” is abbreviated for pre-test/post-test one group (no control group)

^bTypes of stigma measured in the 19 studies are either perceived stigma and/or enacted stigma. Perceived stigma is abbreviated as “PS”; enacted stigma is abbreviated as “ES.”

^c“NS” stands for non-significant results at threshold of .05 for the stigma measure

^dFor Fawole’s [34] study, findings of the stigma item and one other (non-stigma) attitude item were combined to create an overall attitudinal score (not shown)

Table 3

Quality assessment of the 19 studies

1st author	Randomization	Selection bias	Confounding bias	Measurement bias	Intervention focus	Statistics	Overall rating good/fair/poor
RCTs							
Significant							
Apinundecha [25]	Poor/Poor	Good/Poor	Good/Poor	Good/Poor	Good	Good	Good
Gill [27]	Poor/Poor	Poor/Poor	Fair/Poor	Fair/Poor	Fair	Poor	Poor
Klepp [28]	Poor/Poor	Fair/Fair	Good/Poor	Fair/Poor	Fair	Good	Fair
Knaus [29]	Poor/Poor	Poor/Poor	Fair/Poor	Fair/Poor	Poor	Poor	Poor
Krauss [30]	Poor/Poor	Good/Good	Good/Good	Good/Poor	Fair	Good	Good
Not significant							
Ashworth [26]	Fair/Poor	Fair/Good	Fair/Poor	Poor/Poor	Poor	Good	Poor
Markham [31]	Poor/Poor	Poor/Poor	Fair/Poor	Poor/Poor	Fair	Good	Poor
Stewart [32]	Poor/Poor	Poor/Good	Fair/Poor	Fair/Poor	Poor	Poor	Poor
Wu [33]	Poor/Poor	Good/Good	Good/Poor	Poor/Poor	Good	Good	Good
Non-randomized control group (all significant)							
Fawole [34]	Poor/Poor	Good/Good	Good/NA	Poor/Good	Poor	Good	Fair
Kuhn [35]	Poor/Poor	Poor/Fair	Poor/NA	Poor/Poor	Poor	Poor	Poor
Luevswanij [36]	Poor/Poor	Fair/Good	Fair/NA	Poor/Poor	Fair	Poor	Poor
Merakou [37]	Poor/Poor	Poor/Good	Poor/NA	Poor/Poor	Poor	Poor	Poor
Norr [38]	Poor/Poor	Fair/Poor	Fair/NA	Fair/Poor	Fair	Poor	Poor
Yang [24]	Poor/Poor	Poor/Good	Poor/NA	Poor/Poor	Good	Poor	Poor
Without control group							
Significant							
Pisal [39]	NA/NA	Fair/NA	Poor/NA	Poor/NA	Fair	Poor	Poor
Rounds [23]	NA/NA	Fair/NA	Poor/NA	Fair/NA	Poor	Poor	Poor
Sowell [40]	NA/NA	Poor/NA	Poor/NA	Poor/NA	Fair	Poor	Poor
Not significant							
Zachariah [22]	NA/NA	Good/NA	Poor/NA	Fair/NA	Poor	Poor	Poor

N/A Not applicable to rate for this indicator

How to interpret overall ratings

Good Degree of biases, statistics used, and quality of intervention *cannot* refute results

Fair Degree of biases, statistics used, and quality of intervention *may be enough* to refute results

Poor Degree of biases, statistics used, and quality of intervention *are enough* to refute results