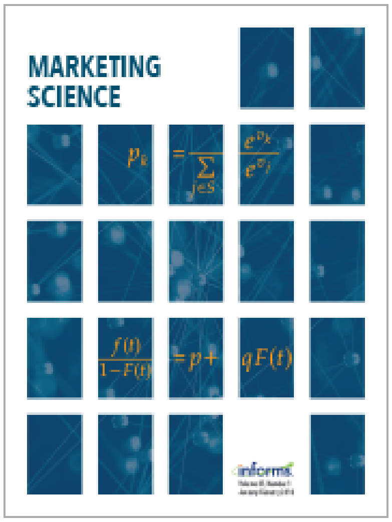


Marketing Science

Publication details, including instructions for authors and subscription information:
http://pubsonline.informs.org

Understanding the Effectiveness of Peer Educator Outreach on Reducing Sexually Transmitted Infections: The Role of Prevention vs. Early Detection

Sam Hui, Parthasarathy Krishnamurthy, Shiv Kumar, Hareesha B. Siddegowda, Prachi Patel

informs
http://pubsonline.informs.org/journal/mksc

MARKETING SCIENCE
Vol. 39, No. 3, May–June 2020, pp. 500–515
ISSN 0732-2399 (print), ISSN 1526-548X (online)

Understanding the Effectiveness of Peer Educator Outreach on Reducing Sexually Transmitted Infections: The Role of Prevention vs. Early Detection

Sam Hui,[a] Parthasarathy Krishnamurthy,[a] Shiv Kumar,[b] Hareesha B. Siddegowda,[c] Prachi Patel[b]

[a] Bauer College of Business, University of Houston, Houston, Texas 77204; [b] Catalyst Management Service, Bengaluru, Karnataka 560094, India; [c] Swasti Health, Pragati Programme, Bengaluru, Karnataka 560094, India

Contact: khui@bauer.uh.edu, http://orcid.org/0000-0002-7846-5794 (SH); partha@uh.edu (PK); shiv@swasti.org (SK); hareesh@swasti.org (HBS); prachi@swasti.org (PP)

Received: December 9, 2016
Revised: January 30, 2018; February 6, 2019
Accepted: March 25, 2019
Published Online in Articles in Advance: August 29, 2019

https://doi.org/10.1287/mksc.2019.1168

Abstract. In an effort to reduce sexually transmitted infections (STI) in developing countries, health organizations often recruit former sex workers as "peer educators" to counsel current sex workers. Although peer educator outreach (PEO) programs have generally been shown to reduce STI, it is not clear whether such efficacy is primarily driven by "prevention" (reducing the infection rate of STI through safe-sex education) or "detection" (educating sex workers about STI symptoms so that they will seek prompt treatment if/when infected). Such differentiation is not only of academic interest, but also has important practical implications on resource management. We develop an integrated Bayesian model to disentangle the role of prevention versus detection in PEO programs. Our results show that PEO programs appear to be not effective in preventing STI, but they do facilitate earlier detection by enhancing sex workers' knowledge and ability to recognize STI symptoms. Simulations based on our model suggest that increasing PEO efforts by 10% from the current level would increase clinic visits by 1.0%, thereby reducing STI prevalence by around 3.0%. Further, we conducted a randomized controlled field experiment that provides directionally consistent evidence that PEO visits are effective in increasing clinic visits among sex workers.

History: K. Sudhir served as the senior editor and Michel Wedel served as associate editor for this article. This paper has been accepted for the *Marketing Science* Special Issue on Health.
Supplemental material: Data are available at https://doi.org/10.1287/mksc.2019.1168.

Keywords: STI • developing country • Bayesian model • policy evaluation • prevention • detection • impact assessment • peer education • field experiment

1. Introduction

The prevalence of sexually transmitted infections (STI) poses a major risk to public health around the world. According to the World Health Organization (WHO), more than one million STI are acquired every day worldwide. In 2015, the number of new infections with the four most common and generally curable STI[1] (chlamydia, gonorrhea, syphilis, and trichomoniasis) is estimated to be around 357 million (WHO 2015). Most of these new infections (estimated to be around 75% to 85%) occur in developing countries, resulting in a major health and financial burden (Mayaud and Mabey 2004). Furthermore, patients in developing countries also suffer from a significantly higher rate of adverse long-term complications because of a generally low level of awareness regarding sexual health and limited knowledge about STI symptoms (Aral et al. 2006), which often leads to delayed and insufficient healthcare seeking.

To combat the spread of STI in developing countries, it is important to prevent and treat STI among sex workers, who are at the highest risk of contracting and transmitting STI to the general public (Parran 1938, World Bank 2008). Thus, many developing countries, for example, India (Basu et al. 2010), Bali (Ford et al. 2000), and the Philippines (Morisky et al. 2006), have deployed peer education outreach (PEO) programs, which employ former sex workers to educate current sex workers (Shahmanesh et al. 2008, WHO 2012) on safe sex/condom use and STI symptoms. Peers are believed to be more effective than professional counselors because they are better at contacting, connecting, empathizing, and communicating with sex workers (Medley et al. 2009). In a typical PEO visit, a peer educator meets sex workers in the field and delivers two specific education "modules": (i) a "prevention" module that focuses on the negotiation with the client regarding sexual practice and the use of condoms and (ii) a "detection" module that focuses on helping the sex worker understand STI symptoms (e.g., vaginal bleeding, rashes) and be her own "first-line triage" so that she will visit the clinic for diagnosis/treatment as soon as she notices any STI symptoms.

Although previous research generally finds that PEO programs are beneficial (Ng et al. 2011, Pickles et al. 2013, Krishnamurthy et al. 2016), to our knowledge prior research has not fully disentangled whether such efficacy is primarily driven by prevention or detection. Understanding the primary drivers of PEO effectiveness is not only conceptually sound because of the way PEO visits are modularized (as discussed earlier), but it also has important managerial and practical implications. For instance, recently, the Bill and Melinda Gates Foundation has started defunding project *Avahan* (a PEO project in India from which we collect our data), and the dedicated community clinics for sex workers have begun to wind down their operations because of a reduction of available funds. If PEO efforts reduce infection rate (prevention), closing dedicated clinics may be less consequential because sex workers are getting infected less often. However, if PEO mainly works through detection, that is, accelerating clinic visits, clinic closure adversely affects access and has a detrimental effect on the spread of STI.

Thus, the goal of this paper is to understand the underlying drivers of PEO efforts: is the effectiveness of PEO efforts mainly a result of prevention (i.e., reducing the primary infection rate of STI) or detection (i.e., allowing sex workers to be more knowledgeable about STI symptoms and, hence, visit the clinic earlier)? Toward that end, we develop an integrated individual-level model, derived from first principles, that jointly captures each sex workers' STI status, clinic-visit behavior, and STI diagnosis over time. Given that a sex worker's STI status over time is only partially observed (i.e., only when the sex worker visits the clinic), we utilize a latent variable approach (Tanner and Wong 1987) with a latent STI state variable for each sex worker for each week and jointly sample it with the rest of the model. Specifically, we model the evolution of a sex worker's STI status over time based on her "baseline" probability of contracting STI each week (which is driven by covariates, such as age, client volume, education level, and place of work); this infection probability can potentially be reduced by the accumulated effect of the PEO prevention module she received, captured through the use of a "adstock"-type variable (e.g., Broadbent 1984 and Dekimpe and Hanssens 2007). Next, conditional on the sex worker's latent STI state, we model the timing by which the sex worker visits the community clinic (if at all) based on her baseline visit propensity and her current STI status, moderated by her knowledge of STI symptoms. Specifically, we assume that a sex worker is more likely to visit the clinic if she has an active STI *and* recognizes the associated symptoms, and her knowledge about STI symptoms is, in turn, driven by the accumulated effect of the PEO detection module she received. Finally, the diagnosis that the sex worker receives at the clinic during a visit reveals her latent STI state at that given time.

Unlike most existing literature that utilizes cross-sectional data sets to evaluate the effectiveness of PEO programs, we utilize an individual-level longitudinal data set of 2,705 adult female sex workers in India from the project *Pragati*. Our data set is comprised of complete longitudinal records of 2,705 sex workers, including peer educator visits, clinic visits, and the associated STI diagnoses. Thus, our data set is considerably richer than Krishnamurthy et al. (2016), who use the same data set, but only consider the *first* clinic visit and do not use any diagnosis data. Our richer data set allows us to identify key model parameters that govern prevention versus detection through two key moments of the data: (i) number of clinic visits and (ii) proportion of positive STI diagnosis given the clinic visit, hence enabling us to, for the first time, disentangle the prevention versus detection drivers of PEO efficacy.

After calibrating our model using Markov Chain Monte Carlo (MCMC), our results suggest that PEO programs appear to be not effective in reducing the infection rate of STI, consistent with Medley et al. (2009) and Thilakavathi et al. (2011), but effective in facilitating earlier detection of STI by enhancing sex workers' sensitivity to STI symptoms. Overall, simulations based on our model suggest that, by increasing PEO efforts by 10% from its current level, the total number of clinic visits is expected to increase by 1.0%, and as a result, STI prevalence (as measured by the total number of active STI person-week) is expected to be reduced by around 3.0%. Further, we find that it is marginally more effective to target PEO efforts to younger sex workers (28 or under), who tend to have higher propensities to contract STI.

Based on these encouraging results obtained from observational data, we conducted a follow-up randomized controlled field experiment in which a separate sample of 147 sex workers were randomly divided into a "treatment" group (who receive a phone-based peer educator "visit") and a "control" group (who do not receive a phone-based peer educator visit). Our experimental results are consistent with our finding that PEO efforts accelerate clinic visits, potentially facilitating accelerated detection and treatment of STI. This provides converging evidence of the efficacy of PEO programs on early detection of STI.

To summarize, the contribution of this paper is fourfold. First, to our knowledge, we collect and utilize the most extensive individual-level longitudinal data set to date to study the efficacy of PEO programs. Our data set goes beyond Krishnamurthy et al. (2016) by studying not only the first clinic visit, but also subsequent visits as well as the associated diagnosis data. Second, we develop a latent variable model to deal with the problem that STI diagnosis is only "partially observed" as it is conditional on clinic visit, a common data limitation in most medical settings. Third, our model allows policy makers to run various counterfactuals to obtain estimates on the "elasticity" of increasing PEO efforts on the number of clinic visits and STI prevalence as well as the effectiveness of targeting specific groups of sex workers based on their demographics (e.g., higher client volume, illiterate, etc.). Finally, we go beyond prior research by conducting a follow-up randomized experiment to address concerns regarding self-selection and endogeneity. Overall, we believe that our results have important managerial implications for healthcare policy makers who would like to further improve the efficacy of PEO programs.

The remainder of this paper is organized as follows. In Section 2, we develop an integrated individual-level model of PEO, STI status, clinic visit, and STI diagnosis. Section 3 describes the individual-level data from project *Pragati* along with key summary statistics and several descriptive analyses. In Section 4, we calibrate the proposed model on the data set and discuss the estimated parameters. We then conduct simulation studies to explore the effectiveness of increasing PEO efforts as well as different targeting strategies in reducing the prevalence of STI. Section 5 describes a randomized controlled field experiment we conducted to rule out self-selection and endogeneity concerns. Finally, Section 6 concludes with directions for future research.

2. An Integrated Model of Peer Educator Outreach, STI Status, Clinic Visit, and STI Diagnosis

In this section, we develop an individual-level integrated model that allows us to disentangle the efficacy of PEO programs into prevention and detection components, corresponding to the two aforementioned modules of a typical PEO program. We outline our latent variable framework and describe our notations in Section 2.1. In Section 2.2, we model the evolution of sex workers' latent STI status over time based on PEO and the treatment they receive from clinic visits. In Section 2.3, we model the incidence of observed clinic visits and STI diagnosis conditional on sex workers' latent STI status as well as the accumulated effect of PEO. Section 2.4 discusses model identification and completes our model specification by briefly describing prior specifications and the MCMC computational procedure.

2.1. A Latent Variable Modeling Framework

Throughout this paper, i indexes sex workers ($i = 1, 2, \ldots, I$), and t indexes time in weeks ($t = 1, 2, \ldots, T_i$), where T_i denotes the total length of weeks that sex worker i is followed in our data set. In addition, $\vec{x}_i$ denotes the various demographics factors for the ith sex worker, which include age, education, average client volume, and place of work (brothel, lodge, street, and home).

The intervention variables from the PEO program are denoted as follows. Let $E_{it} = 1$ if a peer educator met with the ith sex worker on week t, and $E_{it} = 0$ otherwise. As discussed earlier, during that meeting, the peer educator teaches the sex worker about safe sex through the prevention module and educates her on how to recognize common STI symptoms and seek treatment if she notices such symptoms through the detection module.

Next, in terms of the observed outcome variables, we let $v_{it} = 1$ if the sex worker visits the clinic on week t, and $v_{it} = 0$ otherwise. Conditional on the sex worker visiting the clinic on week t (i.e., $v_{it} = 1$), let d_{it} denote the STI diagnosis that the sex worker receives at the clinic. Specifically, we let $d_{it} = 1$ if an STI diagnosis is made, and $d_{it} = 0$ otherwise. As discussed earlier, throughout this paper, we focus on only "curable" bacterial and parasitic STI (most commonly chlamydia, gonorrhea, syphilis, and trichomoniasis; see WHO 2015); we refer to a sex worker infected by curable STI as "STI-active," and a sex worker not infected by curable STI is referred to as "STI-negative" regardless of her non-STI health status.

We now introduce a latent variable framework to link the observed outcome variables (v_{it} and d_{it}) to peer educator visits (E_{it}). For the ith sex worker, we let S_{it} denote her STI status at week t. Given that, based on our data-collection procedure, each sex worker is STI-negative at the beginning of the data-collection period when she registered at the community clinic to the *Pragati* program, we have $S_{i0} = 0$ for each sex worker. The evolution of S_{it} over time depends on the sex worker's baseline propensity of contracting STI, the accumulated effect of PEO prevention education that she has received, and whether she visits the clinic for treatment. We describe our model specification in detail in the next section.

2.2. Model of Latent STI Status over Time

We model the transitions of STI status over time as follows. If a sex worker is STI-active at week t, we assume that, absent treatment at a community clinic, she remains STI-active at week $(t + 1)$. On the other hand, if she visits the clinic on week t, we assume that she would be given appropriate treatment at the clinic and her STI would be cleared by week $(t + 1)$. This model of the evolution of STI status is a simplifying yet reasonable first-order approximation. Medical professionals we spoke to indicate that, on a week-to-week basis, sex workers will remain infected with STI if they are left untreated. In contrast, with appropriate treatment, most common STI can be cleared very quickly, for example, often within a week or several doses of antibiotics in the case of chlamydia infections.[2] Thus, we have

$$[S_{i(t+1)}|S_{it}=1, v_{it}] = \begin{cases} 1 & \text{if } v_{it}=0 \\ 0 & \text{if } v_{it}=1 \end{cases}. \tag{1}$$

Next, if a sex worker is STI-negative at week t ($S_{it}=0$), she may become infected with STI at week $(t+1)$ with infection probability ψ_{it}. That is,

$$\left[S_{i(t+1)}|S_{it}=0, \psi_{it}\right] = \begin{cases} 1 & \text{w.p. } \psi_{\text{it}} \\ 0 & \text{w.p. } 1-\psi_{\text{it}}. \end{cases} \tag{2}$$

The infection probability ψ_{it} is driven by the sex worker's baseline infection propensity ψ_{i0} (which is driven by demographics factors such as age, average client volume, education level, and place of work) as well as the accumulated effect of safe-sex education through the PEO prevention module. Specifically, we model the infection probability ψ_{it} as

$$\text{logit}(\psi_{it}) = \text{logit}(\psi_{i0}) - \theta\left[\sum_{s=1}^{t} E_{is}e^{-\delta(t-s)}\right], \tag{3}$$

where δ is a time-decay factor that captures the decay of the effectiveness of safe-sex education over time because of forgetting. Thus, the term $[\sum_{s=1}^{t} E_{is}e^{-\delta(t-s)}]$ captures the accumulated "stock" of PEO prevention module up to time t in the same spirit as an adstock-type model (Broadbent 1984, Dekimpe and Hanssens 2007).[3] In addition, θ denotes the accumulated effectiveness of the PEO prevention module in reducing the primary infection rate of STI. Next, the baseline propensity for the ith sex worker to contract STI, ψ_{i0}, is modeled as a function of the ith sex worker's demographics as follows:

$$logit(\psi_{i0}) \sim N(\vec{x}_i'\eta_\psi, \sigma_\psi^2), \tag{4}$$

where, as discussed earlier, $\vec{x}_i$ denotes the covariates for the ith sex worker, η_ψ is a vector of regression coefficients that capture the effect of the observed covariates, and σ_ψ^2 captures any remaining unobserved heterogeneity across sex workers.

2.3. Model of Clinic Visits and Diagnosis

Next, we model the incidence of clinic visits and STI diagnosis for each sex worker, conditional on the latent STI status variable. When a sex worker has an active STI ($S_{it}=1$), she is likely to experience one or more STI symptoms and, thus, more likely to visit the community clinic *if she recognizes the specific symptoms*. Her ability to recognize STI symptoms, in turn, depends on her knowledge of and sensitivity to such symptoms, which varies across sex workers and can be enhanced by the PEO prevention module.[4] Thus, we model a sex worker's likelihood of visiting the clinic on week t as a function of her latent baseline propensity to visit the clinic (λ_{it}) and her STI status (S_{it}), together with her knowledge and sensitivity about STI symptoms (β_{it}). We specify

$$\text{logit}(\Pr(v_{it}=1)) = \lambda_{it} + \beta_{it}S_{it}(\beta_{it>0}), \tag{5}$$

where β_{it} represent the sex worker's ability to recognize STI symptoms and λ_{it} denotes the ith sex worker's propensity to visit the community clinic at week t. We model the propensity for clinic visits as follows:

$$\lambda_{it} = \lambda_{i0} + \kappa\left[\sum_{s=1}^{t} v_{is}e^{-\delta(t-s)}\right], \tag{6}$$

where, similar to Equation (3), the term $\kappa[\sum_{s=1}^{t} v_{is}e^{-\delta(t-s)}]$ controls for the effect of the accumulated stock of prior clinic visits up to time t; presumably, the more a sex worker has visited the clinic in the past, the more familiar she is with the clinic and, thus, more likely to visit in the future. Finally, the baseline propensity for clinic visit (λ_{i0}) is modeled as a function of demographics factors as follows:

$$\lambda_{i0} \sim N(x_i'\eta_\lambda, \sigma_\lambda^2), \tag{7}$$

where η_λ is a vector of regression coefficients that capture the effect of the observed covariates and σ_λ^2 captures any remaining unobserved heterogeneity across sex workers.

Next, a sex worker's knowledge and sensitivity about STI symptoms can be enhanced through the detection module of the PEO program, in which the peer educator explains and reminds sex workers of common STI symptoms. We capture the accumulated effect of the PEO detection module by allowing it to drive β_{it} as follows:

$$\log(\beta_{it}) = \log(\beta_{i0}) + \gamma\left[\sum_{s=1}^{t} E_{is}e^{-\delta(t-s)}\right] + \rho\left[\sum_{s=1}^{t} d_{is}e^{-\delta(t-s)}\right], \tag{8}$$

where the term $[\sum_{s=1}^{t} E_{is}e^{-\delta(t-s)}]$ captures the accumulated stock of the PEO detection module that the ith sex worker has received up to time t; thus, γ represents the effectiveness of accumulated PEO efforts on increasing sex workers' knowledge of STI symptoms. The next term on the right, $\rho[\sum_{s=1}^{t} d_{is}e^{-\delta(t-s)}]$, controls for the effect of past STI diagnoses: presumably, a sex worker may become more familiar with STI symptoms after each disease diagnosis she receives at the clinic. Similar to Equations (4) and (7), the baseline knowledge of STI for the ith sex worker, β_{i0}, is modeled as a function of the ith sex worker's demographics as follows:

$$\log(\beta_{i0}) \sim N(x_i'\eta_\beta, \sigma_\beta^2), \tag{9}$$

where η_β is a vector of regression coefficients that capture the effect of the observed covariates and σ_β^2 captures any remaining unobserved heterogeneity.

Finally, given that the ith sex worker visits the community clinic on week t (i.e., $v_{it} = 1$), we assume that her diagnosis at the clinic accurately reveals her STI status (S_{it}); that is,

$$[d_{it}|v_{it} = 1, S_{it}] = S_{it}. \tag{10}$$

Note that the assumption in Equation (10) is fairly reasonable to a first-order approximation given that, in community clinics designated for sex workers in developing countries, STI are often treated "syndromically." That is, sex workers who present with STI symptoms clinically are often treated presumptively under the assumption that they are STI-active (WHO 2007).

Given the model as described in Equations (1)–(10), our main goal is to empirically estimate the key model parameters θ and γ, which govern the effectiveness of PEO programs in terms of prevention and early detection, respectively. In the next section, we discuss model identification and complete our model specification by specifying prior distributions.

2.4. Model Identification, Prior Specification, and Posterior Computation

Given that our model is nonlinear in nature, similar to most biological models in disease modeling and public health research, point identification of model parameters is a major challenge (e.g., Saccomani and Thomaseth 2016). First and foremost, because data on STI diagnosis over time is only partially observed, we have to rely on parametric functional form and distributional assumptions for model identification. As a result, similar to any model that relies on parametric identification, our results are sensitive to the functional form and distributional assumptions in Equations (1)–(10).

Ignoring heterogeneity across sex workers for the moment (which are controlled for by individual-level parameters ψ_{i0} and β_{i0}), we argue conceptually that our key model parameters θ (prevention) and γ (detection) can be plausibly identified through two moments in the data set: (i) the number of clinic visits (y) together with (ii) the proportion of clinic visits that result in a positive STI diagnosis (z). Clearly, the observables (y, z) are a function of (θ, γ); thus, the key model parameters (θ, γ) can be identified through (y, z) if the mapping $(y, z) = f(\theta, \gamma)$ is injective (Lewbel 2019). Based on our model structure and through simulations, we observe the following relationships between the observables and model parameters:

i. The number of clinic visits (y) is strictly monotonically decreasing with respect to θ and strictly monotonically increasing with respect to γ.

ii. The proportion of visits that result in positive STI diagnosis (z) is strictly monotonically decreasing with respect to θ but rather insensitive to γ.

Clearly, increasing θ leads to sex workers being infected less often, hence reducing both the number of clinic visits (because of Equation (5)) and the proportion of visits that result in positive STI diagnosis (as sex workers are less sick to begin with). Increasing γ, on the other hand, has a more nuanced effect. With a more effective detection module, sex workers are more sensitive to STI symptoms when they have them, thus leading to more clinic visits (when the infection rate is held constant) because of Equation (5). Based on our model structure, higher sensitivity to STI symptoms only reduces the "waiting time" between getting infected and going to the clinic; thus, the proportion of visits that result in positive STI diagnosis (z) is only marginally sensitive to changes in γ.

With the observations (i) and (ii), conceptually, one can think of our identification strategy as follows: given that the proportion of visits that result in positive STI diagnosis (z) is insensitive to γ, θ can be identified by z through the strictly monotonic (hence, injective) mapping. Once θ is identified, γ can be identified given θ and the number of clinic visits (y) as the conditional relationship between y and γ is again monotonic (hence, injective). Thus, we argue that our key model parameters can be plausibly identified by the two moments of the data.[5] This identification argument also points to why it is important to consider both clinic visit data and diagnosis data to jointly identify prevention and detection drivers as opposed to Krishnamurthy et al. (2016), who only study the timing of the first clinic visit.

To complete our model specification, the key model parameters that capture prevention and early detection (θ and γ, respectively) are given diffuse, weakly informative $N(0,\ 100^2)$ prior distributions. Similarly, model parameters that control for past clinic visit and past diagnoses, (κ,ρ), are given diffuse, weakly informative $N(0,\ 100^2)$ prior distributions. The time-decay factor $\delta(\delta > 0)$ is given a $N(0,\ 100^2)$ prior distribution that is truncated at zero to ensure that it is positive. Finally, the effects of demographic covariates $(\eta_\psi, \sigma_\psi^2)$, $(\eta_\beta, \sigma_\beta^2)$, and $(\eta_\lambda, \sigma_\lambda^2)$ are each given independent, weakly informative, conjugate prior distributions (Gelman et al. 2003). The joint posterior distribution of all model parameters, along with the latent STI status S_{it}, is sampled using a standard MCMC procedure as described in the appendix. We discard the first 50,000 iterations from the MCMC sampler as a burn-in sample and use the next 100,000 iterations to summarize the posterior distribution (Gelman et al. 2003). The C++ code for the MCMC procedure is available from the authors upon request.

3. Data and Summary Statistics

3.1. Data Overview

Our data set is collected through Project *Pragati* (which means "progress" in the Indian language) in Bengaluru, Karnataka, India. It is an interventional and empowerment program launched in 2005 that aims at controlling STI among female sex workers (Euser et al. 2012). *Pragati* is a joint effort between Swathi Mahila Sangha, a sex worker collective; Swasti, a not-for-profit health resource center; and the *Avahan* initiative of the Bill and Melinda Gates Foundation. A key element within Project *Pragati* is PEO, which aims specifically at reducing the transmissions of STI by educating sex workers about safe sex practices and STI symptoms. As discussed in the introduction, although it is generally believed that peer educator outreach efforts typically lead to reduction of STI over time, it is unclear whether such reduction is driven primarily through prevention or early detection (or both).

Our study sample is comprised of female sex workers who had registered into Project *Pragati* in a community clinic during a four-year period from January 1, 2008, to January 31, 2012, the end of the observation window. We restrict our sample to adult sex workers (age 18 or above) who are STI-negative at the beginning of the data-collection window, resulting in a sample of 2,705 adult female sex workers. To assemble an individual-level data set that describes the detailed history of each sex worker, we merge three separate data sets from *Pragati*'s computerized monitoring and information systems: the registration, outreach, and clinic visit data sets. First, the registration data set provides each female sex worker with a unique ID and records the date she registered into *Pragati*. In addition, it also includes several key demographic variables for each sex worker, including age, average client volume, education level, and place of work (brothel, lodge, street, home). Next, the outreach data set records each peer educator visit to each sex worker. Finally, the clinic visit data set includes the unique ID for each sex worker, her date of clinic visit, and the clinician's diagnosis of STI based on physical examination based on WHO guidelines for syndromic STI management (WHO 2007) as well as the corresponding treatment given. The syndromic STI diagnosis includes vaginitis, cervicitis, pelvic inflammatory disease, urethral discharge, genital ulcer, inguinal swelling, pain and/or scrotal swelling, rectal discharge, anal ulcer, oral lesion, or other oral signs of STI. Of these, the first four account for around 98% of all reported symptoms. Once these three aforementioned data sets are merged, we obtain a timeline for all peer educator visits for each sex worker as well as all clinic visits and corresponding STI diagnoses.

3.2. Summary Statistics and Descriptive Analysis

Table 1 shows several key summary statistics of the data set. As can be seen, sex workers are followed for an average of 140 weeks or a little more than 2.5 years. During the data-observation window, each sex worker visits the community clinic 4.62 times on average with a median of three visits. Roughly 24.5% of sex workers have never visited the clinic during their observation window. About one third of those visits results in an STI diagnosis. The average number of STI diagnoses across a sex worker is about 1.53. The majority (51.2%) of all sex workers are diagnosed with STI at least once during the data-observation window.

Table 1. Summary Statistics of the Data Set

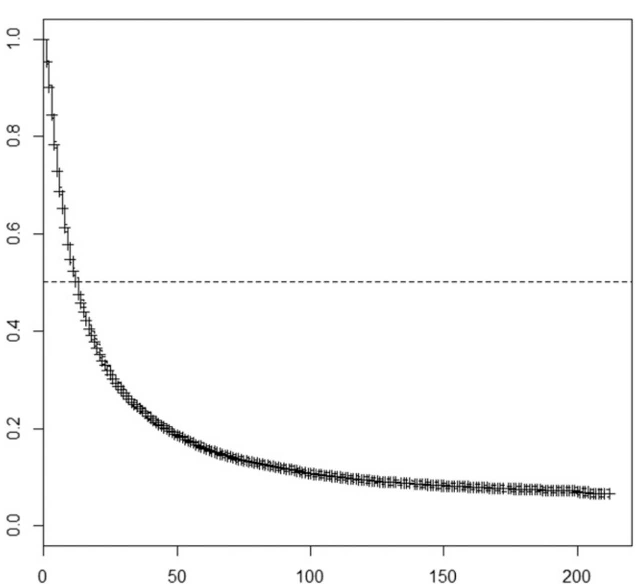
	Mean	Standard deviation	Median	Minimum	Maximum
Length of follow-up period (weeks)	140.02	55.71	162.00	2.00	213.00
Outcome variables					
Number of clinic visits	4.62	5.72	3.00	0.00	58.00
Number of STI diagnosis	1.53	2.27	1.00	0.00	14.00
Intervention variables					
Number of peer educator visits	18.59	17.83	14.00	0.00	88.00
Demographic variables					
Age	29.23	7.49	28.00	18.00	60.00
Average client volume	6.31	4.83	5.00	1.00	30.00
Education (1: illiterate; 0: literate)	44.29%				
Place of work					
Home	67.43%				
Street	20.15%				
Brothel	8.24%				
Lodge	4.18%				

Next, in terms of PEO efforts, the average number of peer educator visits received by sex workers is 18.59 with a median of 14 visits and a range of 0 to 88 visits.[6] This amounts to a total of 50,294 peer educator visits across a total of 378,754 weeks of observation —in other words, a visit every 7.5 weeks on average. The vast majority of sex workers (75%) received at least one peer educator visit during the data-observation window. In terms of demographics, the sex workers in our data set range from 18 to 60 years old with an average age of 29.23. Average client volume is around 6.31 with a range of 1 to 30. Almost half of the sex workers (44.29%) are illiterate and have not received any formal education. Most of the sex workers work at home (67.43%), followed by the street (20.15%), a brothel (8.24%), and a lodge (4.18%).

Next, we perform several descriptive analyses on the intervisit time between clinic visits, one of the key outcome variables that PEO programs are aiming to reduce, using the *survival* library in R. Figure 1 shows the Kaplan–Meier estimate (Kaplan and Meier 1958) of the distribution of intervisit time for clinic visits. As can be seen, across sex workers and clinic visits, the median intervisit time for clinic visits is around 13 weeks (or around three months) with a 95% confidence interval of (12, 13 weeks).

We conduct further descriptive analyses by estimating a mixed-effects logistic regression (Agresti 2013) with the week-by-week incidence of clinic visit (v_{it}) as a dependent variable and the sex worker's demographics covariates (age, education, client volume, and place of work) as well as a variable $[\sum_{s=1}^{t} E_{is}e^{-\delta(t-s)}]$ (as defined in Section 2) that represents the accumulated stock of peer educator visits for each sex worker at any given time as independent variables while controlling for the individual-level differences across sex workers using random effects. In this analysis, we set the value of the decay rate δ to 0.060, the value that is estimated in our full model (as is discussed in Section 4.1). The goal of this analysis is to provide some descriptive evidence on the role of PEO programs in accelerating clinic visits. This mixed-effects logistic regression is estimated using the *lme4* library in R. The results are shown in Table 2.

Figure 1. Kaplan–Meier Survival Curve for Intervisit Time for Clinic Visits



Note. The horizontal broken line indicates the median intervisit time (13 weeks).

Table 2 provides some initial, (relatively) model-free evidence on the effectiveness of peer educator outreach programs on accelerating clinic visits among sex workers. First, the estimated coefficient of the accumulated stock of peer education visits on the incidence of clinic visits is positive and highly significant ($\beta = 0.169; p < .001$), which indicates that a higher intensity of peer educator visits is associated with a higher likelihood of a clinic visit. This is consistent with the previous literature, which shows that peer educator outreach programs increase clinic access (Krishnamurthy et al. 2016). Second, in terms of the role of demographics factors, the estimates in Table 2 suggest that higher client volume ($\beta = -0.011; p < .05$), illiteracy ($\beta = -0.090; p < .05$), and working in a lodge ($\beta = -0.342; p < .01$) are associated with a lower baseline rate of clinic visits. This may suggest specifically targeting PEO efforts to certain demographics groups (e.g., based on education level or place of work) that are less likely to visit the clinic, a strategy that we explore in more detail in Section 4.4.

Though Table 2 provides some initial evidence about the effectiveness of PEO, those descriptive results are difficult to interpret because we have not taken into account the actual STI diagnosis during such a visit (which is itself conditional on a clinic visit occurring). One can argue, for instance, that the

Table 2. Results of Mixed-Effects Logistic Regression with the Incidence of Clinic Visits as Dependent Variable, Sex Worker's Demographics and the Stock of Peer Educator Visits as Fixed Effects, and Sex Worker ID as Random Effects

	Coefficient estimate	*p*-value
Intercept	−4.018***	0.000
Age	−0.003	0.296
Client volume	−0.011*	0.017
Illiterate	−0.090*	0.035
Place of work (brothel)	0.014	0.860
Place of work (lodge)	−0.342**	0.001
Place of work (street)	0.068	0.193
Stock of peer educator visits	0.169***	0.000

***$p < 0.001$; **$p < 0.01$; *$p < 0.05$.

acceleration of clinic visits is partially driven by an *increase* in STI infection rate, thereby causing sex workers to visit the clinic more often as they experience more STI symptoms. Thus, from the descriptive analysis in Table 2 alone, it is unclear how effective PEO efforts are in terms of prevention and early detection; an integrated model is needed to tease apart the differential effects of PEO programs.

3.3. Data Limitations and Caveats

Before presenting the results of our full model, we would like to point out several key limitations and caveats of our data set. As with any observational studies, sex workers' participation in the *Pragati* program is voluntary, which can lead to sample-selection bias (Heckman 1979). For instance, if sex workers who are more receptive to PEO programs are also naturally more likely to engage in safer sex or have higher levels of knowledge about STI symptoms (which would be the case if they are specifically "targeted" by PEO programs), this may spuriously inflate the effectiveness of PEO programs in terms of prevention and detection (respectively). This concern is somewhat alleviated by the fact that we analyze individual-level longitudinal data, with which we control for individual-level heterogeneity through the baseline parameters ψ_{i0} and β_{i0}. Further, we identify the effect of PEO primarily using the *within-subject temporal variations* of PEO exposure, clinic visits, and STI diagnosis, thus minimizing the role of self-selection bias.

To further understand the extent to which targeting might have biased our model results, we conduct informal field interviews with officials at the Swasti organization and peer educators as well as several sex workers. Overall, our field interviews suggest that any targeting efforts tend to be minimal for the following reasons. First, the assignment of sex workers to peer educators (by Swasti) is nontargeted. Each peer educator receives a list of sex workers to counsel along with places to meet them (e.g., home address if the sex worker is home-based, the place of solicitation if the sex work takes place in other venues). According to *Pragati* program officials, no specific targeting efforts have been imposed; thus, the assignment of sex workers to peer educators can be treated as by-and-large random.

Second, peer educators are paid a *flat* monthly honorarium of around 3,000 rupees (equivalent to around \$45 U.S. dollars); their payment is fixed and is independent of the number of meetings with sex workers during the month or the number of clinic visits that they have induced. Thus, from a monetary standpoint, peer educators have no incentive to target specific sex workers.

Third, given the assignment of sex workers to peer educators, the actual meetings between the two parties take place if the two parties "happen to be in the same place at the same time." Specific examples include, for example, a peer educator locating the sex worker at her home, and then the meeting takes place if the sex worker is available; another example is that the peer educator may identify a sex worker at the local bus station or other venues where sex workers solicit clients. These descriptions seem to provide some additional corroborating evidence that both the assignment of sex workers and their meeting mechanisms occur mostly by happenstance and that the role of targeting, if any, does not play a significant role.

Although, admittedly, these descriptions are somewhat vague and nonspecific, which is presumably because the sex trade operates within a gray area of local laws (although sex work itself is not considered illegal, any related activities, for example, solicitation, operating a brothel etc., are illegal in India), on balance, our field interviews tend to suggest that targeting efforts are minimal on the part of the organizing agency and the peer educators.[7] Of course, the only way to fully rule out the potential self-selection issue is to conduct a randomized controlled experiment, in which the treatment of PEO is randomly assigned.[8] Thus, we have conducted a follow-up field experiment study in which sex workers are randomly assigned to treatment and control groups; the results of this field experiment, which are directionally consistent with the main results presented here, are discussed in Section 5.

In addition to the potential issue of self-selection and targeting, as with any other panel data sets, our data may also suffer from the problem of sample attrition (Fitzgerald et al. 1998), for example, a sex worker moves to another city, drops out of sex work, or is being unresponsive to peer educators for any reason during the observational window. Depending on the actual attrition process, this can potentially introduce nonrandom attrition bias. To address this concern, we conduct a robustness check that truncates the data for each sex worker to the time of last contact (either the last peer educator visit or the last visit to a community clinic). Our key results are substantially unchanged and are available upon request.

Despite concerns with respect to self-selection and nonrandom attrition, we would like to stress that our data set is the currently *best available* data set to study the effectiveness of PEO programs in terms of prevention versus detection as it is the only longitudinal data set available at the individual level that ties together peer educator visits, clinic visits, and STI diagnosis.

4. Results

We now present the results of our full model. Section 4.1 discusses key parameter estimates with respect to prevent and detection. Section 4.2 looks at the role of demographics variables on driving baseline propensities among sex workers. In Section 4.3, we estimate the elasticity of clinic visit and STI prevalence with respect to changes in PEO efforts. In Section 4.4, we explore the effectiveness of targeting specific groups of sex workers.

4.1. Effectiveness of Peer Educator Outreach

Table 3 summarizes the posterior distribution of the key model parameters by their posterior means along with 95% posterior intervals. First, in terms of prevention, we find no evidence that safe-sex education through PEO is effective in reducing sex workers' propensities to contract STI ($\theta = 0.004$ with a 95% interval of $(-0.017, 0.025)$). Second, we find that peer educator visits are highly effective in increasing sex workers' sensitivities about STI symptoms ($\gamma = 0.413$); hence, the effectiveness of PEO effort is mainly driven by early detection, which is consistent with the model-free analysis presented in Section 3.2, where we show that PEO accelerates clinic visits. In terms of the control variables (κ and ρ), as expected, we find that prior clinic visits encourage future clinic visits ($\kappa = 0.326$); in addition, we find that past STI diagnoses tend to (modestly) increase sex workers' knowledge about STI symptoms ($\rho = 0.016$). Thus, overall, we find evidence that, although PEO does not seem to help prevent STI, it accelerates the detection and treatment of STI among sex workers by having them visit the community clinic earlier when they are experiencing STI symptoms. As is discussed in Section 4.3, early detection and treatment drastically reduces the prevalence of STI (as measured by the number of active STI person-week) among sex workers.

Furthermore, the decay rate (presumably because of forgetting) of PEO efforts is estimated to be $\delta = 0.060$, a 95% posterior interval of (0.056, 0.065). This corresponds to a "half-life" of roughly 11 weeks with a 95% posterior interval of (10 weeks, 12 weeks). This suggests that the beneficial effects of PEO are fairly long-lasting.

Thus, overall, our results suggest that PEO efforts work mainly through detection but not prevention. This is consistent with the insight we obtain when comparing our full model with two nested models: a benchmark model with prevention only (with γ set to zero) and a benchmark model with detection only (with θ set to zero). The log-Bayes factors, as computed through the Savage–Dickey ratio (Verdinelli and Wasserman 1995) is 6.77 in favor of the full model when compared against the prevention-only model and −264.34 against the full model when compared against the detection-only model. Thus, our results consistently point to the findings that the efficacy of PEO efforts is almost entirely driven by earlier detection. The finding that PEO efforts are effective for detection but not for prevention is consistent with the extant literature on HIV-intervention programs, which shows that peer intervention programs mainly influence behavioral outcomes but not biological outcomes (Medley et al. 2009). Further, our findings corroborate with the findings in Thilakavathi et al. (2011) that find no association between exposure to HIV-prevention programs and odds of infection.

4.2. The Role of Demographics Variables

Next, we explore the association between demographics variables and sex workers' baseline propensities to (i) visit a clinic (λ_{i0}), (ii) contract STI (ψ_{i0}), and (iii) recognize STI symptoms (β_{i0}). The posterior means of the coefficient vectors η_λ, η_ψ, and η_β are shown in Table 4. First, as can be seen in the first column of Table 4, several demographics factors are significantly associated with sex workers' baseline propensities to visit community clinics. Specifically,

Table 3. Parameter Estimates of the Key Model Parameters from the Full Model

Parameter	Interpretation	Posterior mean	95% posterior interval
θ	Prevention: Effectiveness of PEO in reducing chance of contracting STI	0.004	(−0.017, 0.025)
γ	Detection: Effectiveness of PEO in enhancing sex workers' sensitivities to STI symptoms	0.413[a]	(0.380, 0.449)
δ	Decay rate because of forgetting	0.060[a]	(0.056, 0.065)
ρ	Controls for the effect of past STI diagnosis on STI knowledge	0.016[a]	(0.001, 0.043)
κ	Controls for the effect of past clinic visits on clinic visit propensity	0.326[a]	(0.293, 0.358)

[a]95% posterior interval does not cover zero.

Table 4. Parameter Estimates (Role of Demographics) from the Full Model

	λ_{i0} (baseline propensity to visit clinic)	ψ_{i0} (baseline propensity to contract STIs)	β_{i0} (baseline sensitivity to STI symptoms)
Intercept	−4.104[a]	−3.934[a]	−0.792[a]
Age	−0.004	−0.010[a]	−0.001
Client volume	−0.008[b]	−0.011	−0.002
Illiterate	−0.126[a]	0.041	−0.045
Place of work: brothel	0.271[a]	−0.911[a]	0.021
Place of work: lodge	−0.348[a]	−0.277	−0.094
Place of work: street	0.142	−0.087	−0.002

[a]95% posterior interval does not include 0; [b]90% posterior interval does not include zero.

higher client volume, being illiterate, and working at a lodge (compared with working at home) are all associated with a lower baseline rate of clinic visits. Second, turning to sex workers' propensities to contract STI, we see that younger sex workers tend to be at a higher risk of contracting STI. In addition, working at a brothel/lodge is associated with a lower propensity of contracting STI as compared with working at home. Interestingly, no demographics factors are significantly associated with the sex workers' baseline sensitivities to STI symptoms as shown in the third column of Table 4.

Overall, the results in Table 4 suggest that the following demographic subgroups of sex workers are most susceptible to contracting STI and/or also staying infected for a longer period (as they do not often go to the clinic): (i) younger age, (ii) higher client volume, (iii) illiterate, (iv) working at home. This may suggest a targeting strategy in which we focus PEO efforts to certain subgroups based on demographics or place of work. We explore the effectiveness of such targeting strategies in Section 4.4.

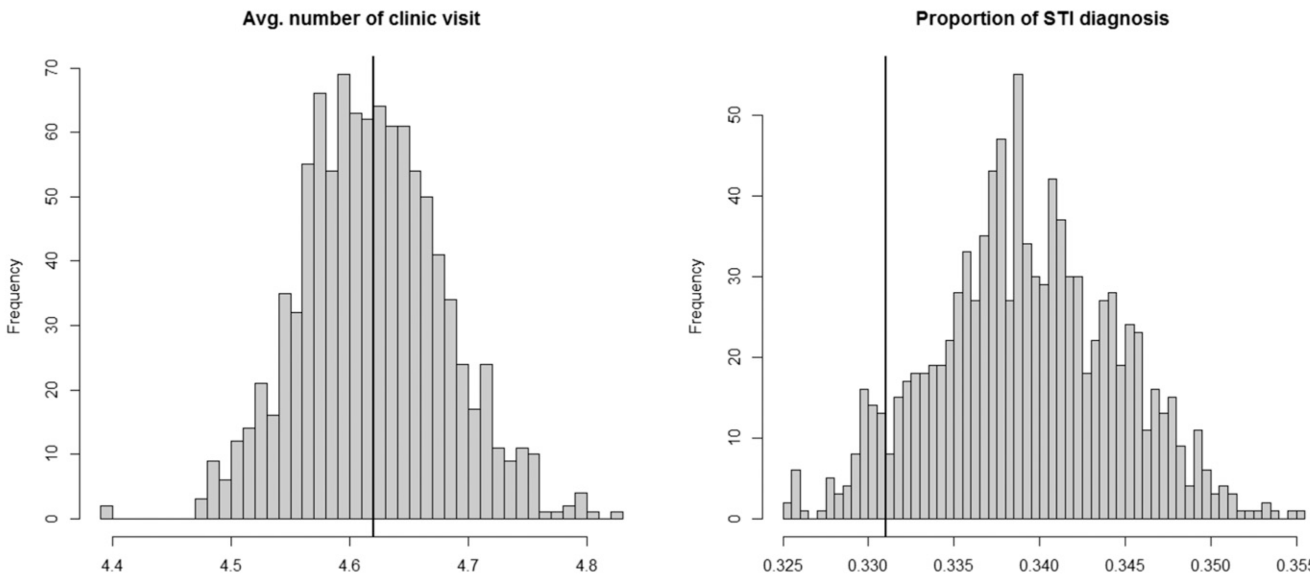
4.3. Elasticity of Clinic Visit Frequency and STI Prevalence with Respect to PEO Efforts

We now estimate the elasticity of clinic visit frequency and STI prevalence (as measured by active STI person-weeks) to the intensity of PEO efforts through a series of simulations. Specifically, starting with the current level of PEO (50,286 total peer educator visits), we vary the level of PEO efforts by increasing the number of peer educator visits by 10% (5,029 more peer educator visits), 20%, (10,058 more visits), and 50% (25,145 more visits). In each case, we randomly assign the additional peer educator visits across all sex worker-weeks that do not currently have a PEO visit (i.e., $E_{it} = 0$) with equal probability. For instance, to increase the number of peer educator visits by 10%, we randomly change 5,029 sex worker-weeks with ($E_{it} = 0$) to ($E_{it} = 1$) and repeat such random assignments 1,000 times. For each repetition, we sample from the posterior distributions of our model parameters from the MCMC sample to simulate each sex worker's STI status as well as her clinic visits and associated diagnosis using the model described in Section 2. We average over the 1,000 random assignments to estimate the average number of clinic visits ($\frac{1}{I}\sum_i\sum_t v_{it}$) and the number of active STI person-weeks ($\sum_i\sum_t S_{it}$), which we use as a measure of STI prevalence across sex workers.

Note that the simulation we conducted with PEO efforts at the current level, that is, using the observed peer educator visits, provides a "posterior predictive check" that allows us to assess model fit (Gelman et al. 1996). By comparing the posterior predictive distribution of key summary metrics (in our case, the average number of clinic visits and the proportion of clinic visits that result in positive STI diagnosis, which, as discussed in Section 2.4, are key for model identification) obtained through posterior simulation and the actual observed values, we can assess how well the model is fitting the data. Specifically, the actual observed data should look "plausible" under the assumed model and parameter estimates (Gelman et al. 2003). The posterior predictive distribution of the average number of clinic visits is shown in Figure 2 with their respective observed values (4.62 and 0.331 as shown in the left and right panels of Table 1, respectively) depicted as vertical lines on the plots. As can be seen, the posterior predicted mean of the average number of clinic visit (4.62) is virtually the same as the actual observed value (4.62) with a posterior predictive p-value (Gelman et al. 2003) of 0.96; similarly, the mean of the posterior predictive distribution for the proportion of clinic visits that results in STI diagnosis (0.337) is close to the actual observed value (0.331) with a posterior predictive p-value of 0.15. Together, they provide some evidence that our model is fairly adequate in describing the data. Note that we cannot compare or assess model fit based on the number of active STI person-weeks as it is unobserved in the actual data.

Simulation results that assess the elasticity of clinic visits and STI prevalence with respect to PEO are summarized in Table 5. Under the current level of PEO efforts, the average number of clinic visits is 4.62, and STI prevalence is estimated to be 75,906 person-weeks. With a 10% increase in PEO efforts, the average number of clinic visits is expected to increase by 1.0% to 4.67; because of earlier detection and treatment of STI, STI prevalence is expected to drop by 3.0% to 73,636 person-weeks. With a 20% increase in PEO efforts, the average number of clinic visits is expected to increase to 4.71 (+2.0% from baseline) with STI prevalence decreasing by 5.8% from baseline. Finally, with a 50% increase in PEO efforts, the average number of clinic visits is expected to increase

Figure 2. Posterior Predictive Distribution of the Average Number of Clinic Visits Across Sex Workers (Left Panel) and Proportion of STI Diagnoses Conditional on Clinic Visit (Right Panel)



Note. The vertical lines indicate the actual observed means (4.62 on the left panel and 0.331 on the right panel).

to 4.85 (+4.9% from baseline), and STI prevalence is expected to decrease by 15.0%. To our knowledge, this is the first time the effect of PEO efforts in terms of increasing clinic visits and reducing STI prevalence has been quantified.

4.4. Effectiveness of Targeting Specific Demographic Groups

The simulations presented in Section 4.3 assign the "additional" peer educator visits randomly across sex worker-weeks (with $E_{it} = 0$) with equal probability. In this section, we explore whether incremental PEO efforts can be further optimized by specifically targeting several high-risk demographics among sex workers. Recall that, in Table 4, we find that younger sex workers are at higher risk of contracting STI, and sex workers with higher client volume, who are illiterate, or who work at home tend to visit the clinic less frequently (hence, if they contract STI, they are less likely to be detected and treated promptly). Thus, we explore the effectiveness of additional PEO efforts that specifically target the following high-risk groups of sex workers:

- Younger age (age ≤28, the median age across sex workers in our data set)
- Higher client volume (volume ≥5, the median reported client volume)
- Illiterate
- Place of work: home

In each of these simulations, we increase PEO efforts by 10% by allocating the additional visits to only the focal demographics. For instance, in the case of targeting sex workers with younger age, we randomly select 5,029 sex worker-weeks only among sex workers with age ≤28 and weeks in which there are currently no peer educator visits ($E_{it} = 0$) and change those to ($E_{it} = 1$). As before, we repeat each simulation 1,000 times and report the average number of clinic visits and STI prevalence across 1,000 runs. The results are summarized in Table 6.

As can be seen, the difference in the average number of clinic visits across different targeting strategies is quite minimal. The changes in STI prevalence are also minimal across the different conditions. Among all these targeting strategies, it appears that targeting younger sex workers results in the largest reduction of

Table 5. Elasticity of the Average Number of Clinic Visits and STI Prevalence with Respect to PEO Efforts

	Average number of clinic visits per sex worker	Percentage change from baseline	Total number of active STIs person-week	Percentage change from baseline
With current PE	4.619	—	75,905.8	—
PE +10%	4.665	+1.00	73,635.8	−2.99
PE +20%	4.713	+2.04	71,503.9	−5.80
PE +50%	4.847	+4.94	64,490.6	−15.04

Note. Number presented is the average across 1,000 sets of simulations.

Table 6. Effectiveness of Strategies That Target PEO Efforts to Specific High-Risk Demographic Groups

	Average number of clinic visits per sex worker	Total number of active STIs person-week	Percentage change from random targeting
Random targeting	4.665	73,635.8	—
Younger age (age ≤28)	4.667	73,455.3	−0.25
High client volume (client volume ≥5)	4.662	73,728.5	+0.13
Illiterate	4.664	73,516.3	−0.16
Place of work: home based	4.670	73,584.8	+0.07

Note. Number presented is the average across 1,000 sets of simulations.

STI prevalence, further reducing the total number of action STI person-weeks to 73,455 (which, compared with the random targeting strategy, represents a −0.25% further reduction). Although the benefit of specific targeting to younger sex workers appears to be marginal, we believe that it is a strategy that is worth considering, especially given anecdotal evidence of widespread underage sex workers in India. One possibility that policy makers may consider is to assign the incremental peer educator visits disproportionally to the highest-risk groups to maximize the effectiveness of PEO efforts, an initiative that the *Pragati* program has been actively investigating in recent years.

5. Follow-up Study: A Randomized Field Experiment

So far, our analysis of longitudinal data suggests that PEO visits facilitate earlier detection and, hence, are associated with more frequent clinic visits. However, as discussed in Section 3.3, our finding can be confounded by self-selection bias as is the case for most observational studies. For this reason, we conduct a pilot field experiment to alleviate self-selection bias and, thus, assess the causal relationship between peer educator outreach and clinic visits. This study was approved by the IRB at the authors' affiliated university.

5.1. Participants, Design, and Method

Between August and September 2017, 147 sex workers were randomly selected from Project *Pragati* and randomly assigned to the treatment condition in which they received one additional peer educator intervention or the control condition in which they did not receive an additional peer educator visit. Initially, we designed the experiment such that the additional peer educator visit that each sex worker in the treatment condition received would be a regular, face-to-face peer educator visit. However, discussions with program officials revealed that the logistical challenges involved in managing a large number of peer educators and having all of them participate in the study are insurmountable given our limited resources. In addition, with multiple peer educators involved, it would be difficult to control for the inherent heterogeneity across peer educators (e.g., some peer educators may be more effective than others, an issue that we briefly return to in Section 6) or their prior interactions with the sex workers (e.g., because sex workers enrolled at different times, some peer educators may already have a long-standing relationship with a sex worker, and others may have a much shorter relationship), thereby introducing significant nonexperimental variations into our study.

Thus, in our pilot field study, the peer educator intervention (the treatment condition) consisted of a *phone call* placed by a single trained peer educator, and the content of the call was aimed specifically at replicating a typical field visit (with both prevention and detection modules). Pragmatically, a phone-based peer educator visit is more managerially tenable in that it does not alter the routine program structure of peer educator field visits but only supplements it at a reasonable cost of time and effort. Further, we only include sex workers who registered relatively recently (who have been registered into the *Pragati* program for less than a year) to ensure that the sex worker has not yet developed a long-standing relationship with her peer educator, and hence, everyone starts with (roughly) the same baseline. In addition, we exclude any sex worker who has received a prior peer educator visit in the 15 days prior to the date of random selection to ensure that their current stock of PEO is relatively low to maximize statistical power of the experiment. Although we recognize that a phone-based peer educator visit is different in nature than a face-to-face peer educator visit, this is the only experimental intervention that is feasible in the current time frame given our resource constraints. In Section 6, we discuss, as a future research direction, a larger-scale field experiment that includes a larger sample size and face-to-face educator visits.

The dependent measure is whether a clinic visit took place in the 40-day period (till the end of the data-collection window) following the start of the experiment. The dependent measure is collected as part of the regular monthly "data sweep" that occurs as part of Project *Pragati*. We analyze the experimental data using intent-to-treat analysis (Hollis and Campbell 1999).

5.2. Experimental Results

The main experimental results are shown in Table 7. As can be seen, 25 out of 71 (35.2%) participants in the treatment condition visited the community clinic at least once during the data-collection window compared with 16 out of 76 (21.1%) participants in the control condition. Comparing across the two conditions, sex workers in the treatment condition are about 67% more likely to visit a community clinic during the data-collection window than sex workers in the control condition. A chi-square test of association shows a marginally significant effect of treatment on odds of clinic visit, $\chi^2_{df=1} = 3.66$, $p < 0.056$ (two-tailed). Thus, overall, the findings from this follow-up study suggest that a phone-based peer educator visit increases the prospect of a clinic visit and facilitates earlier detection of STI. Although the large effect size observed in this study should be taken in the context of the marginal significance of the finding, the result of this field experiment is clearly (directionally) consistent with our main results obtained through observational data.

6. Discussion and Conclusion

In this paper, we disentangle the effectiveness of PEO programs in terms of prevention versus detection using a longitudinal data set from Bengaluru, Karnataka, India. Because STI status is only revealed during clinic visits and, thus, only partially observed, we set up a latent variable framework (Tanner and Wong 1987) and developed an individual-level, integrated model to understand how PEO programs reduce sex workers' risk of contracting STI (prevention) as well as enhancing their knowledge about STI symptoms, hence facilitating earlier detection. We calibrate our model on an individual-level longitudinal data set of 2,705 female sex workers who are registered into Project *Pragati*; to the best of our knowledge, it is, by far, the richest data set that has been used to assess PEO program effectiveness.

Our results suggest that peer educator outreach helps earlier detection and treatment of STI by increasing sex workers' sensitivities and knowledge about STI symptoms. However, PEO programs do not appear to be helpful in preventing STI through safe-sex education not unlike some of the findings in the literature (Medley et al. 2009, Thilakavathi et al. 2011). Overall, if we increase PEO efforts by 10%, we expect the number of clinic visits to increase by 1.0%, thereby reducing STI prevalence (as measured by active STI person-weeks) by around 3.0%. Further, our analysis suggests that certain demographic factors are significantly related to lower baseline clinic visit rate (higher clinic volume, illiterate) and higher baseline risk of contracting STI (younger age, working from home). By specifically targeting younger sex workers, we may further enhance the effectiveness of PEO programs, but the incremental benefit is fairly marginal (as can be seen in Table 6). A follow-up field experiment that we conducted provides additional, directionally consistent evidence that peer educator outreach can be effective in detecting and treating STI earlier.

Even though we find some promising initial evidence for the effectiveness of PEO programs, our research is only the first step toward further understanding and, hence, optimizing peer educator outreach programs, which we defer to future research in this area. Here we provide a list of potential directions for future research.

1. Additional field research: Our finding that safe-sex education does not appear to help prevent STI is intriguing given that safe-sex education has generally been demonstrated to be effective among other population groups (e.g., Kirby 2007). A potential explanation is the power dynamics in the sex worker–client interactions. Using condoms requires the consent of the client, who tends to be the more powerful party in the sex worker–client dyad; for this reason, the sex worker may not be able to engage in prevention behaviors. Thus, the sex worker population may be particularly less susceptible to exhibiting prevention effects (Medley et al. 2009, Thilakavathi et al. 2011). If this power asymmetry is operational, then STI prevention efforts may be more effective when aimed at clients rather sex workers. Alternatively, the prevention mechanism should shift emphasis away from male condoms that require client cooperation to female condoms that permit greater autonomy on the part of the female sex worker. Separately, it is important to assess the impact of STI prevention education on sex workers' STI risk assessments as perceived STI risk often drives actual risky sex behavior (Sychareun et al. 2013).

2. Larger-scale controlled experiments: Given the important caveats and data limitations for our observational study, impact assessment of PEO programs can greatly benefit from a larger-scale, multisite,

Table 7. Summary of Field Experiment Results

	Treatment condition	Control condition
Visited clinic	25 (35.2%)	16 (21.1%)
Did not visit clinic	46 (64.8%)	60 (78.9%)
Total	71	76

randomized control trial. Specifically, sex workers should be randomly assigned to either a "high PEO intensity" or "standard PEO intensity" condition and followed up over time with both clinic-based STI detection as well as field-based, point-of-care STI tests to generate an STI status indication independent of clinic visit propensity.

3. Accounting for heterogeneity across peer educators: In our model, we have assumed that all peer educator visits are equally effective. In reality, there can be significant heterogeneity across peer educators on how effectively they communicate the prevention and detection modules to the sex workers because of personality, social, and demographic differences. Further, the effectiveness of a peer educator visit may depend on how much her background and personality "matches" with that of the sex worker. Future research may look into the heterogeneity across peer educators and identify factors that determine the "success" of a peer educator or to better match peer educators to sex workers to maximize the effectiveness of their interaction.

4. Policy optimization: In the long run, policy makers need to determine what the optimal steady-state level of PEO efforts is. Clearly, the optimal level of PEO efforts depends on the elasticity of incremental PEO efforts on increasing clinic visits and reducing STI prevalence, an estimate that we provide in this paper. In addition, policy makers need to specify not only the per-visit cost of a peer educator visit and the cost of a community clinic visit, but also an estimate of the indirect cost of an active STI person-week among sex workers. The latter estimate should take into account how STI in sex workers are likely to transmit to the population at large, which requires a separate epidemiological model that captures the transmission of STI between sex workers and the general population. Thus, we see our current work as a first step toward a larger trial that culminates in a data-driven decision support system that would help more effectively manage the scare healthcare resources.

Appendix. MCMC Computation Procedure

We briefly describe the MCMC procedure used to calibrate our proposed model. In each iteration of the MCMC draw, we draw from the full conditional distributions of model parameters in the following order: ψ_{i0}, λ_{i0}, β_{i0}, $(\eta_\lambda, \sigma_\lambda^2)$, $(\eta_\psi, \sigma_\psi^2)$, $(\eta_\beta, \sigma_\beta^2)$, κ, (γ, ρ),$(\theta,)$, δ, and finally the latent STI status S_{it}. Each step is outlined as follows.

A.1. Drawing ψ_{i0}, λ_{i0}, and β_{i0}

We use a random walk metropolis algorithm to sample ψ_{i0}. We apply a logit transform on ψ_{i0} and a log transform on λ_{i0}. For each parameter, we use a Gaussian random walk proposal distribution with the mean centered on the value of the previous draw; the variance of the proposal distribution is adjusted to achieve an acceptance rate close to 50%.

A.2. Drawing $(\eta_\lambda, \sigma_\lambda^2)$, $(\eta_\psi, \sigma_\psi^2)$, $(\eta_\beta, \sigma_\beta^2)$

Given that we put a weakly informative, conjugate prior on $(\eta_\lambda, \sigma_\lambda^2)$, given the current draws of λ_{i0}, we can directly draw $(\eta_\lambda, \sigma_\lambda^2)$ using normal–normal conjugate computation (Gelman et al. 2003). Similarly, $(\eta_\psi, \sigma_\psi^2)$ and $(\eta_\beta, \sigma_\beta^2)$ can be sampled through normal–normal conjugate computations given the current draws of ψ_{i0} and β_{i0}, respectively.

A.3. Drawing κ

Because standard conjugate computations are not available to sample κ, we use a random walk metropolis–Hastings algorithm to sample from their posterior distributions. We use a Gaussian random walk proposal distribution with the mean centered on the value of the previous draw; the variance of the proposal distribution is adjusted to achieve an acceptance rate close to 50%.

A.4. Drawing (γ, ρ)

Similarly, because standard conjugate computations are not available to sample (γ, ρ), we use a random walk metropolis–Hastings algorithm to sample from their posterior distributions. Note that we block (γ, ρ) to improve sampling efficiency as an initial run reveals correlation between these two parameters (which is to be expected because both are a part of Equation (8)). Thus, we use a bivariate Gaussian random walk proposal distribution with the mean centered on the value of the previous draw; the variance of the proposal distribution is adjusted to achieve an acceptance rate close to 50% (Gelman et al. 2003).

A.5. Drawing θ

We use a random walk metropolis–Hastings algorithm to sample θ. We use a Gaussian random walk proposal distribution with the mean centered on the value of the previous draw. Again, the variance of the proposal distribution is adjusted to achieve an acceptance rate close to 50%.

A.6. Drawing δ

Likewise, standard conjugate computations are not available to sample δ. We use a random walk metropolis–Hastings algorithm to sample from its posterior distribution. We use a Gaussian random walk proposal distribution with the mean centered on the value of the previous draw. To ensure that δ is positive, we reflect all proposal values off zero (i.e., a draw of −0.2 would become 0.2). Again, the variance of the proposal distribution is adjusted to achieve an acceptance rate close to 50%.

A.7. Drawing S_{it}

Finally, an independent metropolis–Hasting algorithm is used to sample each row of S_{it}. For each sex worker i, (i.e., the ith row of S_{it}), we simulate a "proposal STI history" using the model specified in Equations (3) and (4) given the current draw of all other model parameters. Then, we compute the likelihood of the proposal STI history given Equations (1) and (2), and accept or reject the new draw based on the metropolis–Hastings acceptance probability (Gelman et al. 2003).

As discussed, we discard the first 50,000 iterations from the MCMC sampler as a burn-in sample and use the next 100,000 iterations to summarize the posterior distribution (Gelman et al. 2003). Standard diagnostics confirm that convergence has been reached. The full C++ code and posterior draws are available upon request.

Endnotes

[1] Throughout this paper, we focus only on STI that are curable and, thus, exclude HIV/AIDS and HPV from our consideration. Further, we aggregate all curable STI into a single category, a common practice in the extant literature, for example, Newman et al. (2015) and Ramesh et al. (2010).

[2] See, for example, http://www.cdc.gov/std/chlamydia/stdfact-chlamydia.htm.

[3] We also conduct a robustness check using a Weibull decay function rather an exponential decay. Our results are substantially unchanged and are available upon request.

[4] One may argue that if sex workers already possess generally high levels of knowledge about STI symptoms, any additional education by the PEO prevention module may be ineffective because of a potential "ceiling effect." To address this concern, we conduct a survey of 94 randomly selected sex workers from Project *Pragati*. We present them with a list of 11 symptoms (six STI symptoms: lower abdominal pain, foul-smelling vaginal discharge, burning sensation during urination, genital sore, swelling in the groin, and genital itching with five non-STI symptoms: headache, stress, stomach ulcer, body pain, and vomiting) and ask them to classify each of them as an STI symptom or not. Only 16% of the sex workers correctly classified all 11 symptoms; the average number of correct answer is 7.9 (out of 11), thus ruling out a ceiling effect.

[5] Definitively proving the injectivity of the nonlinear mapping $(y, z) = f(\theta, \gamma)$ (hence, point identification) would requires analytically solving for the inverse mapping $f^{-1}(y, z) = (\theta, \gamma)$, which is impossible. Hence, we provide some further evidence that identification is plausible through a numerical analysis in the spirit of Lagrange et al. (2007) and Walter et al. (2004). Specifically, we simulate (y, z) from a grid of values of (θ, γ) to study the numerical relationship between (y, z) and (θ, γ). From the simulation, the following approximate linear mapping between (y, z) and (θ, γ) can be obtained using regression (with R^2 value of 0.87):

$$y = 3.64 - .75\theta + .63\gamma,$$
$$z = .62 - .15\theta + .00016\gamma.$$

Note that, because of the insensitivity of z on γ, the two equations are linearly independent (the Hessian does not vanish), thus showing that the approximate linear mapping between $(y, z) = f(\theta, \gamma)$ is injective, and hence, (θ, γ) can be identified through (y, z). In addition, we also conduct a "parameter recovery" simulation by using the estimated model parameters (in Section 4) to simulate the time series of sex workers' clinic visits and diagnoses and use the resulting simulated data to estimate the model parameters in the full model. The results show that we can recover the true parameters accurately, adding further corroborating evidence that our model is properly identified.

[6] Note that part of the variation in the number of peer educator visits is driven by the variation in the length of the observation window across sex workers as they are registered into the *Pragati* program at different times. As expected, there is a significant positive correlation between the number of peer educator visits and the length of the observational period ($r = 0.42$, $p < 0.0001$).

[7] To further assess any targeting effects, we also study the relationship between the estimated individual parameters pertaining to infection rate (ψ_{i0}) and sensitivity to STI symptoms (β_{i0}) to the number of received peer educator visits. Both correlations are close to zero and not statistically significant, which corroborates with the program officials' statement that no targeting efforts have been made. We thank an anonymous reviewer for this suggestion.

[8] Alternatively, if targeting is indeed significant, one can use the econometric approach developed by Manchanda et al. (2004) to deal with targeting effects.

Reference

Agresti A (2013) *Categorical Data Analysis*, 3rd ed. (John Wiley and Sons, Hoboken, NJ).

Walter E, Braems I, Jaulin L, Kieffer M (2004) Guaranteed numerical computation as an alternative to computer algebra for testing models for identifiability. Alt R, Frommer A, Kearfott RB, Luther W, eds. *Numerical Software with Result Verification*, Lecture Notes in Computer Science, vol. 2991 (Springer, Berlin, Heidelberg), 124–131.

Aral SO, Over M, Manhart L, Holmes KK (2006) Sexually transmitted infections. Jamison DT, Breman JG, Measham AR, Alleyne G, Claeson M, Evans DB, Jha P, Mills A, Musgrove P, eds. *Disease Control Priorities in Developing Countries*, 2nd ed. (Oxford University Press and the World Bank, New York), 311–330.

Basu I, Jana S, Rotheram-Borus MJ, Swendeman D, Lee S-J, Newman P, Weiss R (2010) HIV prevention among sex workers in India. *J. Acquired Immune Deficiency Syndrome* 36(3):845–852.

Broadbent S (1984) Modeling with adstock. *J. Market Res. Soc.* 26-(October):295–312.

Dekimpe MG, Hanssens DM (2007) Advertising response models. Tellis GJ, Ambler T, eds. *SAGE Handbook of Advertising* (SAGE Publication Ltd., London), P247–263.

Euser S, Souverein D, Gowda P, Gowda C, Grootendorst D, Ramaiah R, Barot S, Kumar S, Jenniskens F, Kumar S, et al. (2012) Pragati: An empowerment programme for female sex workers in Bangalore, India. *Global Health Action* 5:19279.

Fitzgerald J, Gottschalk P, Moffitt R (1998) An analysis of sample attrition in panel data: The Michigan panel study of income dynamics. *J. Human Resources* 33(2):251–299.

Ford K, Wirawan D, Suastina W, Reed B, Muliawan P (2000) Evaluation of a peer education programme for female sex workers in Bali, Indonesia. *Internat. J. STD AIDS*. 11(11): 731–733.

Gelman A, Meng XL, Stern HS (1996) Posterior predictive assessment of model fitness via realized discrepancies (with discussion). *Statistica Sinica* 6(4):733–807.

Gelman A, Carlin JB, Stern HS, Rubin DB (2003) *Bayesian Data Analysis*, 2nd ed. (Chapman & Hall, Boca Raton, FL).

Heckman JJ (1979) Sample selection bias as a specification error. *Econometrica* 47(1):153–161.

Hollis S, Campbell F (1999) What is meant by intention to treat analysis? Survey of published randomized controlled trials. *BMJ* 319(7211):670–674.

Kaplan E, Meier P (1958) Nonparametric estimation from incomplete observation. *J. Amer. Statist. Assoc.* 53(282):457–481.

Kirby D (2007) Sex and HIV programs: Their impact on sexual behavioral of young people throughout the world. *J. Adolescent Health* 40(3):206–217.

Krishnamurthy P, Hui S, Shivkumar N, Gowda C, Pushpalatha R (2016) Assessing the impact of peer educator outreach on the likelihood and acceleration of clinic utilization among sex workers. *PLoS One* 11(7):e0159656.

Lagrange S, Delanoue N, Jaulin L (2007) On sufficient conditions of injectivity: Development of a numerical test algorithm via interval analysis. *Reliable Comput.* 13(5):409–421.

Lewbel A (2019) The identification zoo: Meanings of identification in econometrics. *J. Econom. Literature*. Forthcoming.

Manchanda P, Rossi P, Chintagunta P (2004) Response modeling with nonrandom marketing-mix variables. *J. Marketing Res.* 41(4):467–478.

Mayaud P, Mabey D (2004) Approaches to the control of sexually transmitted infections in developing countries: Old problems and modern challenges. *Sexually Transmitted Infections* 80(3):174–182.

Medley A, Kennedy C, O'Reilly K, Sweat M (2009) Effectiveness of peer education interventions for HIV prevention in developing countries: A systematic review and meta-analysis. *AID Ed. Prevention* 21(3):181–206.

Morisky DE, Stein JA, Chiao C, Ksobiech K, Malow R (2006) Impact of a social influence intervention on condom use and sexually transmitted infections among establishment-based female sex workers in the Philippines: A multilevel analysis. *Health Psych.* 25(5):595–603.

Newman L, Rowley J, Vander Hoorn S, Wijesooriya NS, Unemo M, Low N, Stevens G, Gottlieb S, Kiarie J, Temmerman M (2015) Global estimates of the prevalence and incidence of four curable sexually transmitted infections in 2012 based on systematic review and global reporting. *Plos One* 10(12):e0143304.

Ng M, Gakidou E, Levin-Rector A, Khera A, Murray CJL, Dandona L (2011) Assessment of population-level effect of Avahan, an HIV-prevention initiative in India. *Lancet* 3708(9803):1643–1652.

Parran T (1938) *Shadow on the Land: Syphilis* (Reynal and Hitchcock, New York).

Pickles M, Boily M, Vickerman P, Lowndes CM, Moses S, Blanchard JF, Ramesh BM, Paranjape RS, Alary M (2013) HIV prevention at scale: Has it worked? Evaluation of the impact of the Avahan programme in South India. *Sexually Transmitted Infections* 89(Suppl. 1):A59.

Ramesh BM, Beattie TS, Shajy I, Washington R, Jagannathan L, Reza-Paul S, Blanchard FF, Moses S (2010) Changes in risk behaviours and prevalence of sexually transmitted infections following HIV preventive interventions among female sex workers in five districts in Karnataka state, South India. *Sexually Transmitted Infections* 86(Suppl. 1):17–24.

Saccomani MP, Thomaseth K. (2016) Structural vs. practical identifiability of nonlinear differential equation models in systems biology. Rogato A, Zazzu V, Guarracino M, eds. *Dynamics of Mathematical Models in Biology* (Springer, Cham, Switzerland), 31–41.

Shahmanesh M, Patel V, Mabey D, Cowan F (2008) Effectiveness of interventions for the prevention of HIV and other sexually transmitted infections in female sex workers in resource poor setting: A systematic review. *Tropical Medicine Internat. Health* 13(5):659–679.

Sychareun V, Thomsen S, Chaleunvong K, Faxelid E (2013) Risk perceptions of STIs/HIV and sexual risk behaviours among sexually experienced adolescents in the northern part of Lao PDR. *BMC Public Health* 13:1126.

Tanner MA, Wong WH (1987) The calculation of posterior distributions by data augmentation. *J. Amer. Statist. Assoc.* 82(398): 528–540.

Thilakavathi S, Boopathi K, Girish Kumar CP, Santhakumar A, Senthilkumar R, Eswaramurthy C, Ilaya Bharathy V, Ramakrishnan L, Thongamba G, Adhikary R, et al. (2011) Assessment of the scale, coverage, and outcomes of the Avahan HIV prevention program for female sex workers in Tamil Nadu, India: Is there evidence of an effect? *BMC Public Health* 11(6):S3.

Verdinelli I, Wasserman L (1995) Computing Bayes factors using a generalization of the Savage-Dickey density ratio. *J. Amer. Statist. Assoc.* 90(430):614–618.

World Bank (2008) Sexually transmitted infections in developing countries: Current concepts and strategies on improving STI prevention, treatment, and control. The World Bank Report #42797. Accessed February 4, 2019, http://documents.worldbank.org/curated/en/2008/03/9068926/sexually-transmitted-infections-developing-countries-current-concepts-strategies-improving-sti-prevention-treatment-control.

World Health Organization (2007) *Introducing STI Syndromic Case Management: Training Modules for the Syndromic Management of Sexually Transmitted Infections*, 2nd ed. (World Health Organization, Geneva).

World Health Organization (2012), Prevention and treatment of HIV and other sexually transmitted infections for sex workers in low- and middle-income countries: Recommendation for a public health approach. Accessed February 4, 2019, http://www.who.int/hiv/pub/guidelines/sex_worker/en/

World Health Organization (2015) Sexually transmitted infections (STIs). World Health Organization fact sheet #110. Accessed February 4, 2019, http://www.who.int/mediacentre/factsheets/fs110/en/.