

1 **An Incognito Standardized Patient Approach for Measuring and Reducing Intersectional** 2 **Healthcare Stigma**

3
4 **Authors:** M. Kumi Smith^{1,§}, Danyang Luo², Siyan Meng³, Yunqing Fei⁴, Wei Zhang⁵, Joseph Tucker^{6,7},
5 Chongyi Wei³, Weiming Tang⁶, Ligang Yang⁸, Benny L Joyner⁹, Shujie Huang⁸, Cheng Wang⁸, Bin
6 Yang⁸ and Sean Y Sylvia¹⁰.

7
8
9 ¹Division of Epidemiology & Community Health, University of Minnesota Twin Cities, Minneapolis,
10 MN, USA; ²Zhitong Guangzhou LGBT Center, Guangzhou, China; ³Department of Health Behavior,
11 Society and Policy, Rutgers University, New Brunswick, NY, USA; ⁴Center for Global Health and Social
12 Responsibility, University of Minnesota Twin Cities, Minneapolis, MN, USA; ⁵The University of North
13 Carolina at Chapel Hill UNC Project-China, Guangzhou, China; ⁶Institute for Global Health and
14 Infectious Diseases, School of Medicine, University of North Carolina, Chapel Hill, NC, USA; ⁷London
15 School of Hygiene and Tropical Medicine, London, UK; ⁸Dermatology Hospital of Southern Medical
16 University, Guangzhou, China; ⁹Departments of Pediatrics, Anesthesia and Social Medicine, University
17 of North Carolina, Chapel Hill, NC, USA; ¹⁰Department of Health Policy & Management, University of
18 North Carolina, Chapel Hill, NC, USA

19
20 [§]Corresponding author: M. Kumi Smith, 1300 South 2n Street, Suite 300, Minneapolis, MN 55454, USA.
21 Tel: +1 (612) 301-3051. (smi00831@umn.edu)

22
23
24 Email Addresses of all authors:

25
26 MKS: smi00831@umn.edu; DL: ryan.luo5@gmail.com; SM: kmsiyan@gmail.com; YF:
27 feixx042@umn.edu; WZ: krystalwz19@yahoo.com; JT: jdtucker@med.unc.edu; CW:
28 cw788@sph.rutgers.edu; WT: weiming_tang@med.unc.edu; LY: yanglg3@hotmail.com; BLJ:
29 benny@unc.edu; SH: huangshj_jm@126.com; CW: wangcheng090705@gmail.com; BY:
30 yangbin101@hotmail.com; SYS: sean_sylvia@unc.edu.

31 **Keywords:**

32 Standardized Patient, Healthcare Stigma, LGBT, HIV/AIDS, China,
33
34

35 **ABSTRACT**

36

37 **Background:** Consistent evidence highlights the role of stigma in impairing healthcare access in people
38 living with HIV (PLWH), men who have sex with men (MSM), and people with both identities. We
39 developed an incognito standardized patient (SP) approach to obtain observations of providers to inform a
40 tailored, relevant, and culturally appropriate stigma reduction training. Our pilot cluster randomized
41 control trial assessed the feasibility, acceptability, and preliminary effects of an intervention to reduce
42 HIV stigma, anti-gay stigma, and intersectional stigma.

43

44 **Methods:** Design of the intervention was informed by the results of a baseline round of incognito visits in
45 which SPs presented standardized cases to consenting doctors. The HIV status and sexual orientation of
46 each case was randomly varied, and stigma was quantified as differences in care across scenarios. Care
47 quality was measured in terms of diagnostic testing, diagnostic effort, and patient-centered care. Impact of
48 the training, which consisted of didactic, experiential, and discussion-based modules, was assessed by
49 analyzing results of a follow-up round of SP visits using linear fixed effects regression models.

50

51 **Results:** Feasibility and acceptability among the 55 provider participants was high. We had a 87.3%
52 recruitment rate and 74.5% completion rate of planned visits (N=238) with no adverse events. Every
53 participant found the training content “highly useful” or “useful.” Preliminary effects suggest that,
54 relative to the referent case (HIV negative straight man), the intervention positively impacted testing for
55 HIV negative MSM (0.05 percentage points [PP], 95% CI, -0.24, 0.33) and diagnostic effort in HIV
56 positive MSM (0.23 standard deviation [SD] improvement, 95% CI, -0.92, 1.37). Patient-centered care
57 only improved for HIV positive straight cases post-training relative to the referent group (SD, 0.57; 95%
58 CI, -0.39, 1.53). All estimates lacked statistical precision, an expected outcome of a pilot RCT.

59

60 **Conclusions:** Our pilot RCT demonstrated high feasibility, acceptability, and several areas of impact for
61 an intervention to reduce enacted healthcare stigma in a low-/middle-income country setting. The
62 relatively lower impact of our intervention on care outcomes for PLWH suggests that future trainings
63 should include more clinical content to boost provider confidence in the safe and respectful management
64 of patients with HIV.

65

66

67 INTRODUCTION

68

69 Stigma is one of the most commonly cited reasons for poor uptake of evidence-based interventions (EBI)
70 for HIV prevention such as testing or preexposure prophylaxis.[1] Stigma impedes care access through
71 several mechanisms. Patients who feel undeserving of care (internalized stigma) or who fear
72 discrimination (anticipated stigma), for example, may abstain from seeking EBI, whereas clinical staff
73 may discriminate by offering suboptimal care or turning patients away (enacted stigma). Uptake is lowest
74 in key populations such as gay and bisexual men who have sex with men (MSM) who face additional
75 stigma on account of social taboos against same-sex or other behaviors. The interaction of multiple
76 stigmatized identities, or intersectional stigma, is an increasingly recognized aspect of HIV stigma, but
77 few effective interventions exist.[2–6] [5,7] A better understanding of the layered nature of stigma is
78 necessary for disentangling its effects and for informing design of effective interventions.[2,7,8]

79

80 Interventions on stigma, including those in healthcare settings, have been plentiful enough to motivate
81 multiple reviews on the topic.[9–16] All the reviews note the abundance of impactful interventions, but
82 caution that pervasive issues with design quality and methodological rigor limit meaningful insights as to
83 what actually works. Central among the methodological issues is that of stigma measurement, itself a
84 long-standing topic of discussion.[16–19] Enacted stigma in particular is singled out for its inherent
85 challenges,[9,16] namely how few providers are likely to admit to discrimination and the difficulty of
86 surveying patients who may lack the clinical knowledge to objectively evaluate their quality of care.
87 Compounding these challenges is that of how to best measure intersectional stigma, as victims of this type
88 of stigma cannot easily disentangle discrimination originating from different sources.[20]

89

90 Our team developed a stigma reduction training for providers by creating a novel approach to measuring
91 enacted, intersectional stigma experienced by MSM, people living with HIV (PLWH), and people with
92 both identities. We measure stigma by deploying incognito standardized patients (SP) or trained actors
93 who present standardized disease cases for the purposes of clinical observation. Providers consent to
94 visits in advance but are not told when they will take place, allowing insights into their true behaviors in
95 real clinical settings. By randomly varying the sexual orientation and HIV status of presented cases, we
96 can quantify stigma as the difference in care quality received across scenarios. Results are shared with
97 separate community advisory boards (CAB) of local providers and MSM to solicit their views on stigma
98 drivers, allowing us to enlist the insights of people closest to the problem.[21] We hypothesize that our
99 intervention can more effectively reduce stigma by giving trainees more tailored and hands on content
100 than traditional curricula informed by theoretical reasoning alone.[22–24] Given the importance of STI

101 care as an entry point into the HIV prevention continuum for MSM at highest risk of HIV infection, we
102 worked with providers of sexually transmitted infection (STI) clinics.

103
104 Results of the baseline round of stigma assessment using incognito SP visits are reported in a separate
105 manuscript. Here we present the results of a pilot randomized control trial (RCT) to assess the feasibility,
106 acceptability, and preliminary impact of a stigma reduction intervention informed by an unannounced SP
107 approach to measuring stigma against MSM, PLWH, and people with both identities.

108
109

110 **METHODS**

111

112 The pilot cluster RCT was conducted between March 2021 and August 2022. In brief, a round of
113 unannounced SP visits were conducted at baseline, results of which were used to inform the design of the
114 stigma reduction intervention. Following delivery of the training intervention, a second round of
115 unannounced SP visits were conducted to assess the key outcomes of feasibility, acceptability, and
116 preliminary impact of the intervention. An overview of the study and data collection procedures is
117 provided in Figure 1.

118

119 **STUDY SETTING**

120

121 This pilot RCT was conducted in Guangzhou, China. Guangzhou, a city of over 11 million residents,
122 embodies the hallmarks of the Chinese HIV epidemic in MSM: rapidly rising HIV and STI prevalence
123 and prevalent healthcare stigma.[25,26] The prefectural municipality of Guangzhou is made up of 11
124 urban districts, 10 from which we recruited our study clinics (Figure 2). All field activities were
125 conducted in partnership with the Dermatology Hospital of the Southern Medical University (SMU)
126 which oversees surveillance, clinical practice, and implementation of disease control policy through a
127 province-wide network of >400 STI practices. Practices included both standalone clinics and specialty
128 wards within larger hospitals.

129

130 **RECRUITMENT & RANDOMIZATION**

131

132 Our sampling frame consisted of STI practices listed on SMU's network roster located in Guangzhou.
133 The first stage of our two-step recruitment process consisted of approaching clinic or ward directors in
134 person to explain the study goals and if they agreed, to obtain a list of providers employed at their

135 practice. Eligible practices were those with formal government medical accreditation and with the
136 capacity to provide enzyme-linked immunosorbent assay testing for HIV, treponemal (e.g., *Treponema*
137 *pallidum* particle agglutination) and non-treponemal tests (e.g. rapid plasma reagin] for syphilis. In the
138 second stage of recruitment we approached practice providers individually to inform them of study goals,
139 answer questions, and obtain consent. Eligible providers were 1) at least 18 years of age; 2) certified to
140 provide STI related care in Guangdong province, and 3) planning to remain at the practice clinic for at
141 least one year.

142
143 Randomization at the practice level used a modified Zelen design in which control arm participants are
144 not informed that they are part of an RCT.[27–29] This approach, which has been applied in fields
145 ranging from STIs to chronic disease, seeks to minimize bias from potential compensatory behaviors of
146 participants who are knowingly assigned to the control arm (i.e. the John Henry effect).[30] During
147 consent procedures, all providers, regardless of arm, were instructed to document details of any suspected
148 SP visits which could then be verified post study to assess the rate of SP detection by providers. All
149 providers then took part in a 15-minute survey administered by trained study staff to provide information
150 on their demographics and professional background. A facility level survey was also completed by
151 appropriate clinic staff to document clinic characteristics such as staff size and patient load.

152 153 DATA COLLECTION

154
155 We conducted two waves of incognito SP visits with consenting providers: one set at baseline and a
156 second set four months after the intervention training. The presented case was a young male, age 20-40,
157 presenting with complaints of primary syphilis (i.e. a recently healed chancre on the penis) and
158 condomless sex. We chose primary syphilis because it allowed for plausible presentation by healthy
159 volunteers and because of the public health significance of timely treatment for MSM and PLWH, both of
160 whom experience elevated incidence relative to other populations. We randomly varied the HIV status
161 and sexual orientation with by visit to obtain care quality measures on each of four scenarios: HIV
162 negative straight man, HIV negative MSM, HIV positive straight man, and HIV positive MSM (scenarios
163 are hereafter referred to as “referent,” “MSM only,” “HIV only”, and “intersectional). SPs announced the
164 HIV status and sexual orientation at the top of each visit using scripted opening lines.

165
166 SP hiring and training were conducted in close collaboration with the Zhitong LGBT Center, a
167 Guangzhou-based community-based organization (CBO) specializing in LGBT+ advocacy and health
168 promotion. Candidate SPs who met the basic descriptions of the role took part in a two day training which

169 sought to achieve 1) realistic and consistent case presentations across all SPs and 2) a consensus on
170 interpretation of items on the healthcare quality checklist used for data collection. Training activities
171 included a review of study materials (scripts, checklists, safety protocol), role plays, and field testing. A
172 second fresher training was held prior to the follow-up round of visits.

173
174 Provider participants received three SP visits per wave. The following case features were randomized by
175 facility and within each facility: the specific SP conducting the visit, the case scenarios presented, and the
176 order in which each provider received the case scenarios. SP visits with the same provider were spaced
177 out by a minimum of two weeks to reduce risk of SP detection. Immediately after each visit,
178 accompanying study staff conducted data collection with SP using the healthcare quality checklist and a
179 brief qualitative interview to capture visit features that might have been missed by the checklist.
180 Throughout the study, SPs met periodically as a team to discuss checklists and ensure mutually consistent
181 interpretations of items and ratings.

182 183 MEASURES

184
185 Data collection was conducted using a healthcare quality checklist designed to capture multiple
186 dimensions of provider behavior that could theoretically shape clinical and interpersonal patient
187 experiences (see **Supplemental materials**). Clinical items were informed by national diagnosis and
188 treatment guidelines on syphilis case management,[31] and interpersonal items by input from both of our
189 CABs. Data collected using the healthcare quality checklist was then operationalized into stigma
190 measures by estimating differences in the quality of care between each of the test cases and the referent
191 case (straight, HIV negative). Care quality was measured across the three domains of care including
192 syphilis testing, diagnostic effort, and patient-centered care, details of which are included in the
193 **Supplemental materials**.

194
195 We also measured feasibility of the intervention as recruitment rates, retention rates, and incidence of
196 adverse events. Acceptability was measured using responses to a self-administered online survey
197 distributed to providers following final study visits. Respondents were asked to evaluate various
198 components on a 4-point Likert scale ranging from “dissatisfied” to “very satisfied.”

199 200 INTERVENTION DEVELOPMENT

201
202 Details of enacted healthcare stigma observed at baseline are reported in a separate manuscript currently

203 under review. Briefly, baseline visits documented evidence of all three forms of stigma: HIV stigma,
204 MSM stigma (or heterosexism), and intersectional stigma. Stigma was most apparent in the lower quality
205 of clinical care (syphilis testing, diagnostic effort) received by the various scenarios, whereas patient-
206 centered care scores were similar across all four scenarios. The study team, with input from CAB
207 members, preliminarily concluded that stigma towards gay and HIV positive patients manifested in this
208 setting as neglect or avoidance, most likely due to providers' lack of knowledge or exposure to these
209 types of patients.

210
211 Analysis of our baseline results and input from our two CABs were evaluated using the information,
212 motivation, and behavioral skills (IMB) model of behavior change as a guiding framework.[32] The
213 resulting intervention centered on the three following goals: 1) to convey the significance of the STI
214 epidemic in marginalized populations including MSM and PLWH; 2) to persuade providers of the public
215 health significance of their role; and 3) to strengthen their skills in communicating with marginalized
216 patients. Components included didactic sessions supported by prepared animated videos, role plays with
217 volunteer SPs uninvolved in study visits, and group discussions on strategies to improve patient-provider
218 communication. To accommodate provider schedules, a fully online version of the training was offered to
219 providers who could not attend the in-person event.

220

221 *Analysis*

222

223 Preliminary effects of the pilot intervention were calculated using an intent-to-treat (ITT) approach. Our
224 primary outcomes of MSM stigma, HIV stigma, and intersectional stigma were conceptualized as the
225 difference in care quality received in each test scenario (i.e. MSM only, HIV only, intersectional) and the
226 referent scenario (i.e. HIV negative straight man). Preliminary intervention effects were calculated for
227 each domain of care: syphilis testing, diagnostic effort, patient-centered care. Linear ordinary least
228 squares models were used to estimate training effects for each of the primary outcomes as follows:

$$229 \quad Y_{idj} = \alpha + (\beta_1 * Training_j) + \sum_{\substack{c \in (MSM, HIV-); \\ (Straight, HIV+); \\ (MSM, HIV+)}} (\beta_2 * Scenario_c) + (\delta_c * Training_j * Scenario_c) + \bar{Y}_{(t-1)dj} + \varepsilon_{idj}$$

230

231 where Y_{idj} is a given post-intervention outcome for a clinical encounter between an SP with doctor d in
232 facility j ; $training$ is a binary indicator for study arm assignment (1 for facilities randomized to the
233 training intervention and 0 for control facilities); and $Scenario_c$ is a set of binary indicators to designate
234 the presented test scenario. To enhance statistical precision, each model controlled for the average of the

235 outcome in all interactions with doctor d at baseline, $\bar{Y}_{(t-1)dj}$. As a cluster RCT, randomization was
236 conducted by facility; thus our heteroskedasticity-consistent standard errors adjusted for clustering at the
237 facility level.[33,34]

238
239 This regression specification yields estimates for δ_c , three primary quantities of interest for each outcome.
240 As increases in each outcome represent better care, this measures the extent to which differences in care
241 relative to the referent group changed due to the intervention. Estimates above the null value of 0
242 represent a reduction in enacted stigma. Estimates below the null correspond to a smaller improvement in
243 care quality relative to the referent scenario and are interpreted as *increases* in a particular form of stigma.

244
245 *Ethical Approvals*

246
247 This study was approved by the institutional review boards at the University of Minnesota, University of
248 North Carolina, and the Dermatology Hospital of the Southern Medical University. All study participants
249 provided written informed consent in Chinese. Our study followed the Consolidated Standards of
250 Reporting Trials (CONSORT) reporting guideline.

251

252

253 **RESULTS**

254

255 *Sample Characteristics & Follow-up*

256

257 The CONSORT flow diagram (Figure 3)[35] shows the recruitment and retention patterns for the study.
258 According to the Zelen design, randomization occurred before providers were approached for individual
259 consent. Following randomization, all eligible providers in each clinic were approached by study staff. Of
260 the 34 who were approached at intervention arm clinics, 30 agreed to participate (88.2%); and of the 29 in
261 the control arm, 25 (86.2%) agreed.

262

263 Providers had a mean age of 42 (standard deviation [SD], 9), were mostly male (62%), and were mostly
264 assistant- or intermediate level clinicians (63%, as opposed to associate- or senior-level clinicians; Table
265 1). Enrolled clinics had a mean patient load of 861 weekly outpatients (SD 579) and employed an average
266 of 4.8 clinicians (SD 2.9) and 2.2 (SD 2.6) support staff (Table 2).

267

268 A baseline round of 123 unannounced SP visits were successfully completed: 72 in the intervention arm,

269 51 in the control arm. Of the 165 total planned visits, 123 (74.5%) were completed, for an 80%
270 completion rate in the intervention arm and 68% in the control arm. Four months following intervention
271 completion, we completed 115 (69.9%) of the 165 planned second wave visits: 71.1% in the intervention
272 arm and 68% in the control arm. Reasons for non-completion included providers being on temporary
273 leave (e.g. medical, maternity), leaving their position at the clinic, or being unavailable for visits after two
274 attempts. No adverse events were reported during any of the visits. 41 (87.2%) of providers did not
275 suspect or did not know if they had received an SP visit. None of the remaining 6 were able to recall visit
276 dates, precluding our ability to verify SP detection.

277

278 *Feasibility & Acceptability*

279

280 Regarding feasibility, 77.3% of the 22 clinic directors we approached agreed to participate in our study
281 (N=17), and 87.3% of the 66 individual providers approached enrolled in the study (N=55). 14 of the 30
282 participants in the treatment arm took part in the intervention (46.7%), 6 in the in-person training and 8 in
283 the synchronous online intervention. The remainder (41.7%) received intervention materials via Wechat
284 (a popular text messaging app). The most commonly reported reasons for not attending the intervention
285 included time conflicts, not having enough time, and facing unexpected COVID related travel restrictions.

286

287 Regarding acceptability of the intervention, all who took part in either in-person or online training
288 reported that each training component (didactic, role play, group discussion) was “very useful” or “quite
289 useful” (as opposed to “a little bit useful” and “not useful”). The portion that reported content as “very
290 useful” was higher for the in-person attendees (66.7-83.3%) than for online attendees (37.5%). Similar
291 patterns were observed in reported rates of satisfaction with aspects of the training delivery including
292 pacing, difficulty, and quality of material, as well as the knowledge and preparation of trainers.

293

294 *Preliminary effect*

295

296 Estimates of the marginal intervention effects on each type of stigma with each of the three domains of
297 care are shown in Figure 4. In terms of syphilis testing, the intervention had a modest positive impact on
298 MSM stigma (0.05 percentage points [PP]; 95% confidence interval [CI], -0.24, 0.33) and negatively
299 impacted HIV and intersectional stigmas (-0.23 PP, 95% CI, -0.55, 0.085 and -0.07 PP; 95% CI, -0.38,
300 0.25, respectively), though all of these estimates lacked statistical precision. In terms of diagnostic effort,
301 the intervention had negative impacts on MSM and HIV stigma (SD, -0.45, 95% CI, -1.60, 0.69; SD,
302 -0.97, 95% CI, -1.83, -0.11, respectively) and a positive impact on intersectional stigma (SD, 0.23; 95%

303 CI, -0.92, 1.37), though only the estimate for HIV stigma was statistically significant. Lastly for patient-
304 centered care, we observed near null effects for MSM and intersectional stigma (SD, -0.01, 95% CI, -
305 0.83, 0.82; SD, 0.00; 95% CI, -1.15, 1.14) and a positive impact on HIV stigma (SD, 0.57; 95% CI, -
306 0.39, 1.53), though once more all estimates lacked statistical precision.

307
308 Additional insights are provided by stigma-specific estimates which quantify the absolute (vs. relative)
309 impact of the intervention impact on each domain of care (Figure 5). These results indicate that the
310 intervention had an absolute positive impact on the probability of syphilis testing for HIV negative MSM,
311 the amount of diagnostic effort invested in HIV positive MSM, and the patient-centeredness of care for
312 HIV positive MSM, though all of these estimates lacked statistical precision.

313

314

315 **DISCUSSION**

316

317 This pilot RCT documented high feasibility, acceptability, and several areas of impact for an intervention
318 to reduce enacted healthcare stigma in a low/middle-income setting. The incognito SP approach provided
319 unique insights into the particular ways that HIV stigma, MSM stigma, and intersectional stigma manifest
320 in clinical settings, facilitating the creation of an intervention more responsive to providers' actual service
321 gaps and training needs. The objectivity of the incognito SP measure also allowed for a more rigorous
322 evaluation of program impact. Our findings build off the one other known application of the incognito SP
323 approach to measure enacted healthcare stigma, in which Li et al. dispatched SPs to compare behaviors of
324 providers assigned to treatment versus control arms of an HIV stigma training in China.[36]

325

326 A central feature of our intervention was its distinct impacts on cases of different sexual orientations. That
327 is, it appeared to improve clinical care—i.e. syphilis testing, diagnostic effort—for MSM of either HIV
328 status but not for straight PLWH. This may be partially due to the mixing of our intervention message
329 with those of our collaborators at SMU who as provincial STI authorities regularly emphasize the
330 importance of MSM-facing clinical care to combat the regional syphilis epidemic. In addition, the siloed
331 nature of China's STI and HIV care systems may mean that our providers are far less likely to encounter
332 patients who are PLWH than MSM. Due to their low exposure to PLWH—and lack of HIV related
333 training—non-HIV specialists may therefore resort to avoidance or needless referral in their rare
334 encounters with an HIV patient.[37–39] However the fact that treatment arm providers' patient-centered
335 care scores improved for HIV positive scenarios is an encouraging sign of their general receptiveness for
336 ways to improve care for PLWH.

337

338 Though preliminary, results of this pilot RCT provide valuable guidance for future interventions and
339 stigma research. First, the siloed nature of care for PLWH, an initially useful strategy to rapidly roll-out
340 HIV treatment in many LMIC, may have also inadvertently deprived non-specialists of much needed
341 training and experience in managing PLWH. This worsens stigma when non-specialists who feel
342 underprepared resort to the understandable but problematic habit of perfunctory visits or needless referral.
343 Future trainings should therefore feature relevant clinical skills in the correct, safe, and respectful
344 management of patients with HIV. Second, our study exposed the challenges of targeting multiple forms
345 of stigma, particularly when one identity (same sex behaviors among men) is a clinically valid risk factor
346 for the other (HIV infection). Though some amount of trait-based generalization has a role in good
347 clinical practice[40] excessive profiling can strain patient-provider relations.[41–43] Future trainings
348 must therefore navigate the balance between healthy and harmful use of clinically salient patient history
349 to improve care quality for key populations.

350

351 Findings from this pilot RCT should be interpreted in light of several key limitations. First, though our
352 study was not powered to detect intervention effects, statistical power could have been strengthened by
353 better participation in the intervention. Suboptimal participating was due in part to COVID-19 related
354 prevention and reporting duties which consumed much of the limited free time our participants had. Many
355 were restricted from travel due to COVID-related lockdowns. Future interventions may address these
356 issues by dispatching academic details to deliver intervention content at each clinic. Second, our measure
357 of patient-centered care relied on subjective assessments by individual SPs. Our SP training included
358 team exercises to align their interpretation of items and rating scales across team members, but future uses
359 of the unannounced SP approach may benefit from additional booster trainings to improve inter-reliability
360 of SP reporting in order to improve data validity.

361

362 This study provides valuable proof of concept for the safe use of incognito SP visits for assessing,
363 developing, and evaluating effective interventions to reduce enacted healthcare stigma in Chinese STI
364 settings. The incognito SP approach is particularly well suited to measuring the more subtle and indirect
365 forms of enacted stigma which may be less perceptible to individual patients, rendering patient surveys
366 less reliable. The approach is also highly adaptable for capturing stigma from other sources (e.g.
367 race/ethnicity, gender, sex, age) and a potentially powerful to assess the impact of structural
368 stigma[44,45] by including facility-level features (e.g. support staff attitudes, clinic practices) as part of
369 SP data collection. Finally, SP approaches create opportunities for meaningful co-creation of
370 interventions with both stigmatized communities (MSM) and the intervention targets (providers),

371 furthering the principles of community based participatory research.[46]

372

373

374 COMPETING INTERESTS

375 The authors declare that they have no competing interests.

376

377 AUTHOR CONTRIBUTIONS

378 M.K.S. and S.Y.S. conceived the study. M.K.S., S.Y.S., D.L, S.M, J.T, W.T, L.Y, B.L.J, S.H, C.W and

379 B,Y collaborated in the design of the methods and intervention. M.K.S. and S.Y.S. conducted the

380 analyses with help from D.L and S.M. M.K.S wrote the initial drafts. S.Y.S provided inputs for the

381 revision, J.T, W.T provide comments for the draft. All authors read and approved the final manuscript.

382

383 ACKNOWLEDGEMENTS

384 We thank all hospitals and providers for taking part in this study, as well as the The Guangdong STD

385 Control Center. We also thank the Zhitong LGBT center for their support and connection to their

386 community members. We would also like to express our gratitude for the hard work of all the SPs.

387

388 FUNDING

389 This research was supported by NIH grants R34MH121251 (MKS and SYS)

390

391 **Trials Registration:** NCT04896216

392

393

394

395

396 **REFERENCES**

397

- 398 1. Sullivan MC, Rosen AO, Allen A, Benbella D, Camacho G, Cortopassi AC, et al. Falling Short of
399 the First 90: HIV Stigma and HIV Testing Research in the 90–90–90 Era. *AIDS Behav.* 2020 Feb
400 1;24(2):357–62.
- 401 2. Friedland BA, Sprague L, Nyblade LC, Baral SD, Pulerwitz J, Gottert A, et al. Measuring
402 intersecting stigma among key populations living with HIV: implementing the people living with
403 HIV Stigma Index 2.0. *JAIDS J Acquir Immune Defic Syndr.* 2018;21(Supp 5):e25131.
- 404 3. Rausch D. Promoting Reductions in Intersectional StigMa (PRISM) to Improve the HIV Prevention
405 Continuum. 2018.
- 406 4. Rice WS, Logie CH, Napoles TM, Walcott M, Batchelder AW, Kempf MC, et al. Perceptions of
407 intersectional stigma among diverse women living with HIV in the United States. *Soc Sci Med.*
408 2018;
- 409 5. Logie CH, James LI, Tharao W, Loutfy MR. HIV, gender, race, sexual orientation, and sex work: A
410 qualitative study of intersectional stigma experienced by HIV-positive women in Ontario, Canada.
411 *PLoS Med.* 2011;8(11).
- 412 6. Goodin BR, Owens MA, White DM, Strath LJ, Gonzalez C, Rainey RL, et al. Intersectional health-
413 related stigma in persons living with HIV and chronic pain: implications for depressive symptoms.
414 *AIDS Care - Psychol Socio-Med Asp AIDSHIV.* 2018;
- 415 7. Nyblade LC. Measuring HIV stigma: Existing knowledge and gaps. *Psychol Health Med.*
416 2006;11(3):335–45.
- 417 8. Reidpath DD, Chan KY. A method for the quantitative analysis of the layering of HIV-related
418 stigma. *AIDS Care.* 2005;17(4):425–32.
- 419 9. Smith MK, Xu RH, Hunt SL, Wei C, Tucker JD, Tang W, et al. Combating HIV stigma in low- and
420 middle-income healthcare settings: a scoping review. *J Int AIDS Soc.* 2020;23(8):e25553.
- 421 10. Nyblade L, Stangl A, Weiss E, Ashburn K. Combating HIV stigma in health care settings: What
422 works? *J Int AIDS Soc.* 2009;12(1).

- 423 11. Sengupta S, Banks B, Jonas D, Miles MS, Smith GC. HIV interventions to reduce HIV/AIDS
424 stigma: A systematic review. *AIDS Behav.* 2011;15(6):1075–87.
- 425 12. Feyissa GT, Lockwood C, Woldie M, Munn Z. Reducing HIV-related stigma and discrimination in
426 healthcare settings: A systematic review of quantitative evidence. *PLOS ONE.* 2019 Jan
427 25;14(1):e0211298.
- 428 13. Brown L, Macintyre K, Trujillo L. Interventions to reduce HIV/AIDS stigma: what have we
429 learned? *AIDS Educ Prev.* 2003/03/12 ed. 2003;15(1):49–69.
- 430 14. Mak WWS, Mo PKH, Ma GYK, Lam MYY. Meta-analysis and systematic review of studies on the
431 effectiveness of HIV stigma reduction programs. *Soc Sci Med.* 2017;188:30–40.
- 432 15. Stangl AL, Lloyd JK, Brady LM, Holland CE, Baral SD. A systematic review of interventions to
433 reduce HIV-related stigma and discrimination from 2002 to 2013: how far have we come? *J Int*
434 *AIDS Soc.* 2013;16(3 Suppl 2).
- 435 16. Nyblade L, MacQuarrie K. Can We Measure HIV/AIDS-Related Stigma and Discrimination?
436 Current Knowledge about Quantifying Stigma in Developing Countries. *Usaid.* 2006;(January):1–
437 21.
- 438 17. Earnshaw VA, Chaudoir SR. From conceptualizing to measuring HIV stigma: A review of HIV
439 stigma mechanism measures. *AIDS Behav.* 2009;13(6):1160–77.
- 440 18. Karver TS, Atkins K, Fonner VA, Rodriguez-Diaz CE, Sweat MD, Taggart T, et al. HIV-Related
441 Intersectional Stigma and Discrimination Measurement: State of the Science. *Am J Public Health.*
442 2022;112(S4):S420–32.
- 443 19. Mahendra VS, Gilborn L, Bharat S, Mudoi RJ, Gupta I, George B, et al. Understanding and
444 measuring AIDS-related stigma in health care settings: A developing country perspective. *Sahara J.*
445 2007 Aug;4(2):616–25.
- 446 20. Harnois CE, Bastos JL. The promise and pitfalls of intersectional scale development. *Soc Sci Med.*
447 2019 Feb;223:73–6.
- 448 21. Li L, Wu Z, Liang LJ, Lin C, Guan J, Jia M, et al. Reducing HIV-related stigma in health care
449 settings: a randomized controlled trial in China. *Am J Public Health.* 2013;103(2):286–92.

- 450 22. Sengupta S, Banks B, Jonas D, Miles MS, Smith GC. HIV interventions to reduce HIV/AIDS
451 stigma: A systematic review. *AIDS Behav.* 2011;15(6):1075–87.
- 452 23. Mahajan APA, Sayles JN, Patel V a, Remien RH, Szekeres G, Coates TJ. Sigma in the HIV/AIDS
453 epidemic: A review of the literature and recommendations for the way forward. *Aids.*
454 2008;22(Suppl 2):S67–79.
- 455 24. Nyblade L, Stangl A, Weiss E, Ashburn K. Combating HIV stigma in health care settings: What
456 works? *J Int AIDS Soc.* 2009;12(1).
- 457 25. Cao B, Zhao P, Bien C, Pan S, Tang W, Watson J, et al. Linking young men who have sex with men
458 (YMSM) to STI physicians: A nationwide cross-sectional survey in China. *BMC Infect Dis.*
459 2018;18(1):228.
- 460 26. Tang W, Mao J, Tang S, Liu C, Mollan K, Cao B, et al. Disclosure of sexual orientation to health
461 professionals in China: Results from an online cross-sectional study. *J Int AIDS Soc.* 2017;20(1):1–
462 9.
- 463 27. Zelen M. A New Design for Randomized Clinical Trials. *N Engl J Med.* 1979 May;300(22):1242–5.
- 464 28. Homer CSE, E HCS. Using the Zelen design in randomized controlled trials: debates and
465 controversies. *J Adv Nurs.* 2002 Apr;38(2):200–7.
- 466 29. Torgerson DJ, Roland M. What is Zelen’s design? *BMJ.* 1998 Feb;316(7131):606–606.
- 467 30. Schellings R, Kessels AG, ter Riet G, Knottnerus JA, Sturmans F. Randomized consent designs in
468 randomized controlled trials: Systematic literature search. *Contemp Clin Trials.* 2006;27(4):320–32.
- 469 31. Chen ZQ, Zhang GC, Gong XD, Lin C, Gao X, Liang GJ, et al. Syphilis in China: results of a
470 national surveillance programme. *The Lancet.* 2007 Jan;369(9556):132–8.
- 471 32. Fisher JD, Fisher WA, Williams SS, Malloy TE. Empirical tests of an information-motivation-
472 behavioral skills model of AIDS-preventive behavior with gay men and heterosexual university
473 students. *Health Psychol.* 1994;13(3):238–50.
- 474 33. Lin W. Agnostic notes on regression adjustments to experimental data: Reexamining Freedman’s
475 critique. *Ann Appl Stat.* 2013 Mar;7(1):295–318.

- 476 34. MacKinnon JG, White H. Some heteroskedasticity-consistent covariance matrix estimators with
477 improved finite sample properties. *J Econom.* 1985 Sep 1;29(3):305–25.
- 478 35. Begg C, Cho M, Eastwood S, Horton R, Moher D, Olkin I, et al. Improving the Quality of
479 Reporting of Randomized Controlled Trials: The CONSORT Statement. *JAMA.* 1996 Aug
480 28;276(8):637–9.
- 481 36. Li L, Lin C, Guan J. Using standardized patients to evaluate hospital-based intervention outcomes.
482 *Int J Epidemiol.* 2014 Jun 1;43(3):897–903.
- 483 37. Li L, Wu Z, Wu S, Zhao Y, Jia M, Yan Z. HIV-related stigma in health care settings: a survey of
484 service providers in China. *AIDS Patient Care STDs.* 2007 Oct;21(10):753–62.
- 485 38. Li L, Liang L, Lin C, Wu Z, Wen Y. Individual attitudes and perceived social norms: Reports on
486 HIV/AIDS-related stigma among service providers in China: *International Journal of Psychology:*
487 *Vol 44, No 6. Int J Psychol [Internet].* 2009 [cited 2023 Jun 6];44(6). Available from:
488 <https://www.tandfonline.com/doi/abs/10.1080/00207590802644774>
- 489 39. Cai G, Moji K, Honda S, Wu X, Zhang K. Inequality and Unwillingness to Care for People Living
490 with HIV/AIDS: A Survey of Medical Professionals in Southeast China. *AIDS Patient Care STDs*
491 *[Internet].* 2007 [cited 2023 Jun 6];21(8). Available from:
492 <https://www.liebertpub.com/doi/epdf/10.1089/apc.2006.0162>
- 493 40. Wasserman D. Is Racial Profiling More Benign in Medicine Than Law Enforcement? *J Ethics.*
494 2011;15(1–2):119–29.
- 495 41. Burgess DJ, Warren J, Phelan S, Dovidio J, van Ryn M. Stereotype Threat and Health Disparities:
496 What Medical Educators and Future Physicians Need to Know. *J Gen Intern Med.*
497 2010;25(S2):169–77.
- 498 42. Betancourt JR, Green AR, Emilio Carrillo J, Park ER. Cultural competence and health care
499 disparities: Key perspectives and trends. *Health Aff (Millwood).* 2005;24(2):499–505.
- 500 43. Harper GW. Sex isn't that simple: culture and context in HIV prevention interventions for gay and
501 bisexual male adolescents. *Am Psychol.* 2007;62(8):803–19.
- 502 44. Rao D, Elshafei A, Nguyen M, Hatzenbuehler ML, Frey S, Go VF. A systematic review of multi-
503 level stigma interventions: State of the science and future directions. *BMC Med.* 2019;17(1):1–11.

- 504 45. Nyblade L, Mbuya-Brown RJ, Ezekiel MJ, Addo NA, Sabasaba AN, Atuahene K, et al. A total
505 facility approach to reducing HIV stigma in health facilities: implementation process and lessons
506 learned. *AIDS Lond Engl*. 2020;34(October 2019):S93–102.
- 507 46. Ayala G, Sprague L, Leigh-Ann van der Merwe L, Thomas RM, Chang J, Arreola S, et al. Peer-
508 And community-led responses to HIV: A scoping review. *PLoS ONE*. 2021;16(12 December).
- 509
510

TABLES

Table 1. Provider characteristics participating in the study in Guangzhou, China.

	Total	Control	Treatment	Difference ¹	95% CI ¹	p-value ¹
Total (N)	55	25	30			
Provider Age, Mean (SD)	42 (9)	43 (9)	40 (9)	2.9	-1.9, 7.8	0.2
Gender, n (%)				0.07	-0.46, 0.60	
Female	21 (38.2)	10 (40)	11 (36.7)			
Male	34 (61.8)	15 (60)	19 (63.3)			
Education, n (%)				0.43	-0.11, 0.97	
Professional School	2 (3.6)	0 (0)	2 (6.7)			
Bachelor Degree	13 (23.6)	5 (20)	8 (26.7)			
Graduate Degree	40 (72.7)	20 (80)	20 (66.7)			
Title, n (%)				0.65	0.11, 1.2	
Assistant-level clinician	10 (18.2)	3 (12)	7 (23.3)			
Intermediate-level clinician	25 (45.4)	12 (48)	13 (43.3)			
Associate-level clinician	16 (29)	7 (28)	9 (30)			
Senior clinician	3 (5.5)	3 (12)	0 (0)			
Other	1 (1.8)	0 (0)	1 (3.3)			
Average Work Hours per Week, Mean (SD)	39.6 (5.4)	39.0 (4.0)	40.0 (6.4)	-0.96	-3.8, 1.9	0.5
Average Patient Load per Week, Mean (SD)	49 (20)	41 (18)	55 (20)	-14	-24, -3.7	0.009

¹ Welch Two Sample t-test; Standardized Mean Difference

CI: Confidence Interval

Table 2. Clinic characteristics.

	Total	Control	Treatment	Difference¹	95% CI¹	p-value¹
Outpatients in past week, Mean (SD)	861 (579)	766 (614)	944 (570)	-178	-795, 439	0.5
Number of clinicians, Mean (SD)	4.9 (2.9)	4.88 (3.7)	4.9 (2.2)	-0.01	-3.3, 3.3	>0.9
Number of clinicians above associate level, Mean (SD)	2.1 (2.1)	2.25 (2.6)	2.0 (1.7)	0.25	-2.1, 2.6	0.8
Number of support staff, Mean (SD)	2.2 (2.6)	1.63 (1.8)	2.7 (3.1)	-1	-3.7, 1.6	0.4
Clinic provides treatment for occupational HIV exposure, n (%)	16 (94%)	8 (100%)	8 (89%)	11%		0.6
Clinic provides clinician training on patient-centered care, n (%)	14 (88%)	6 (75%)	8 (100%)	-25%		0.3
Unknown	1	0	1			

1. Welch Two Sample t-test; 3-sample test for equality of proportions without continuity correction

CI: Confidence Interval

FIGURE CAPTIONS

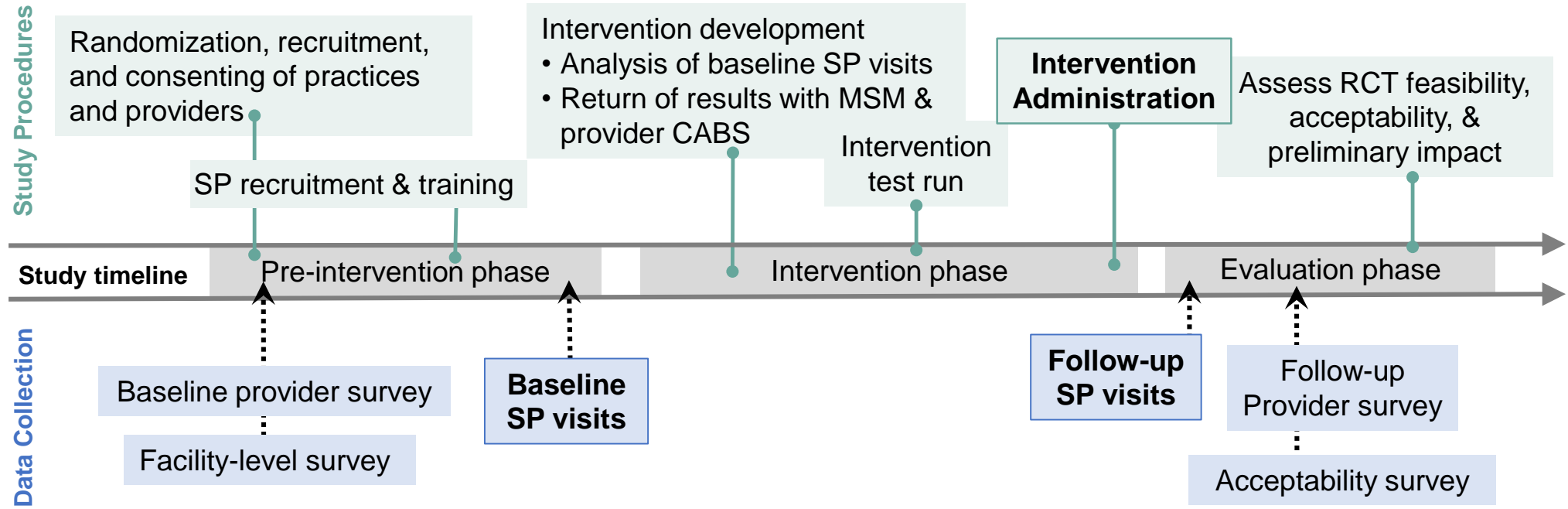
Figure 1. Study timeline, divided into procedures and data collection procedures.

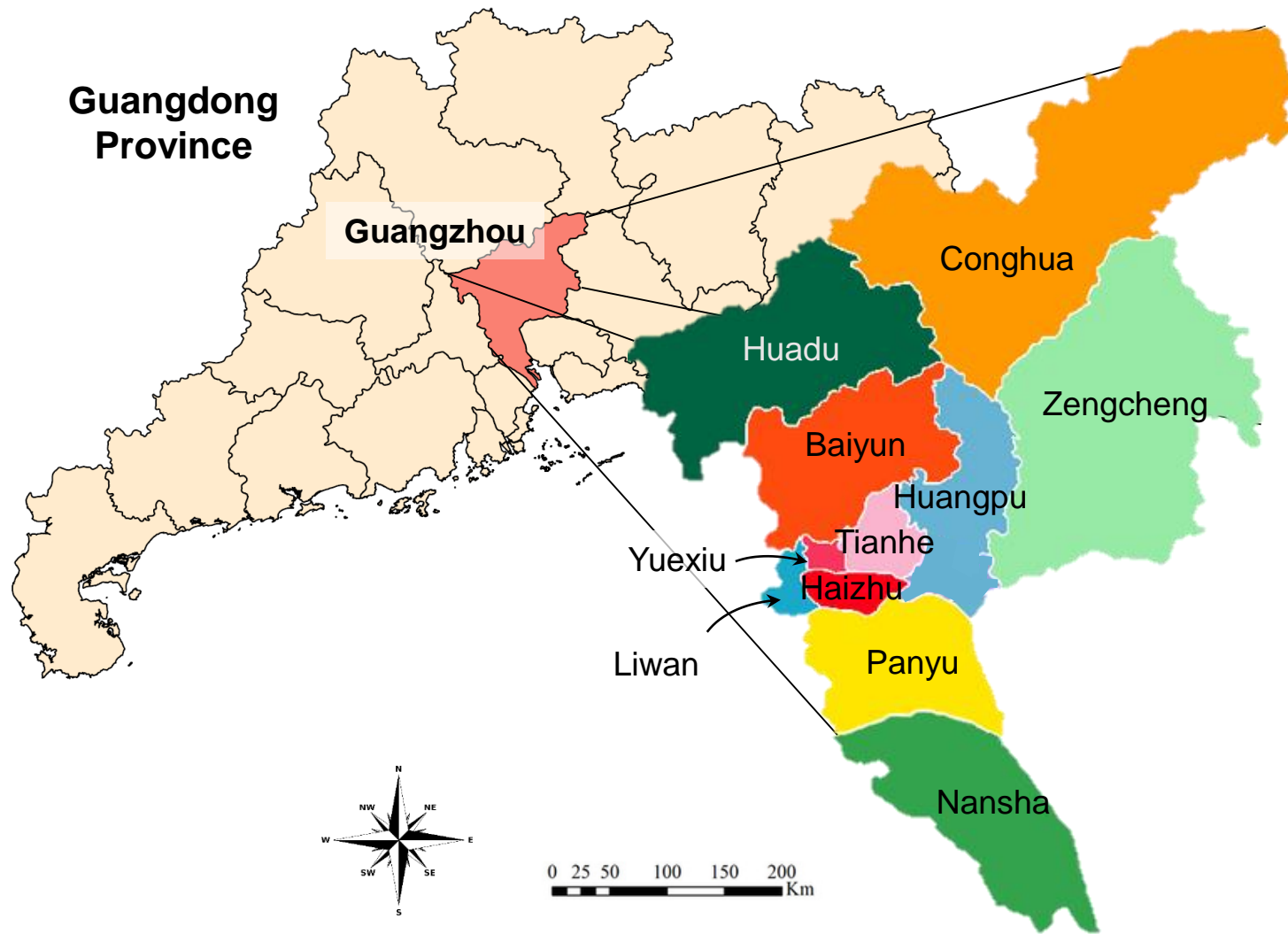
Figure 2. Map of the study setting. The prefectural municipality of Guangzhou is made up of 11 urban districts.

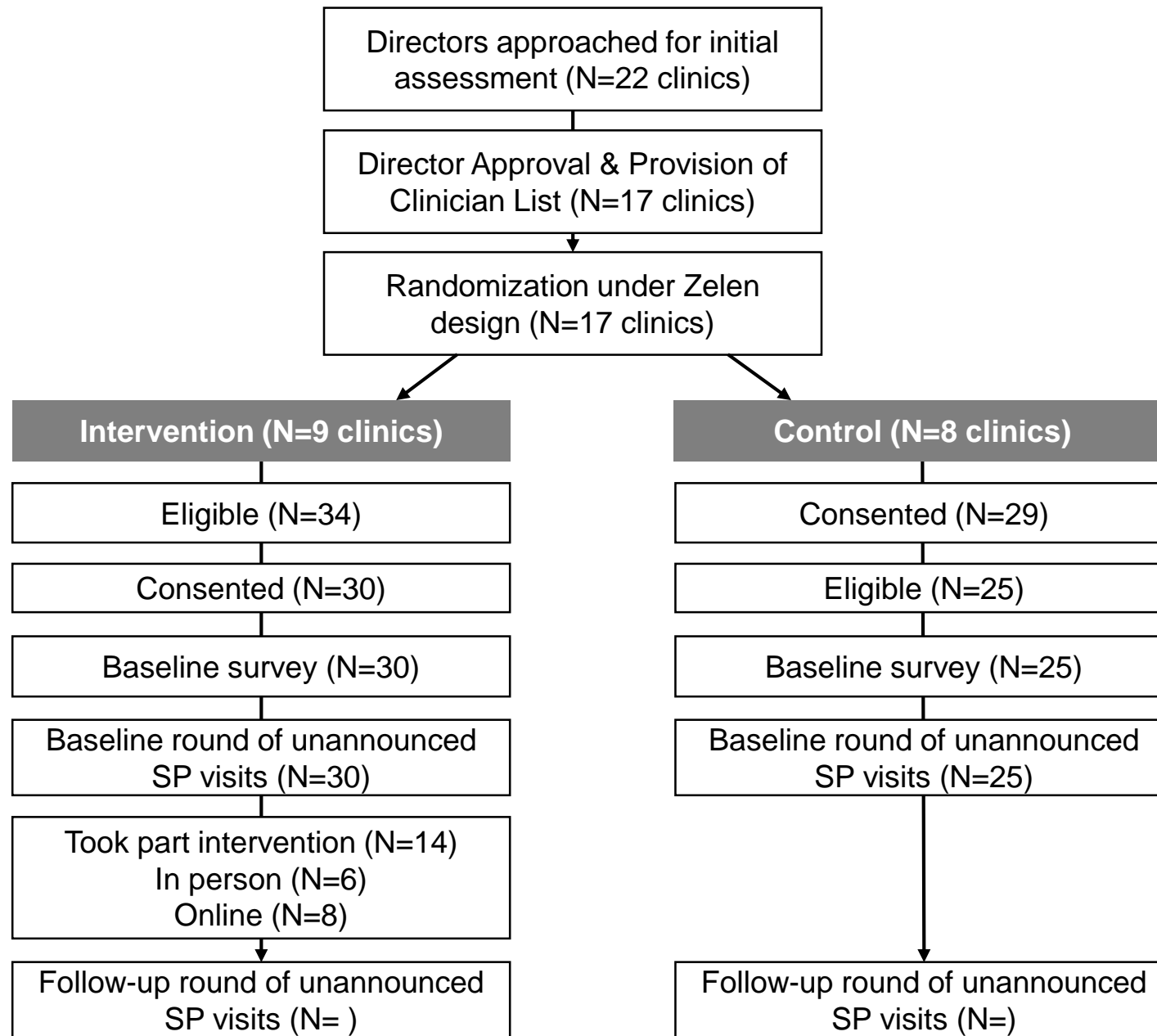
Figure 3. CONSORT map of the pilot cluster randomized control trial to reduce enacted healthcare stigma.

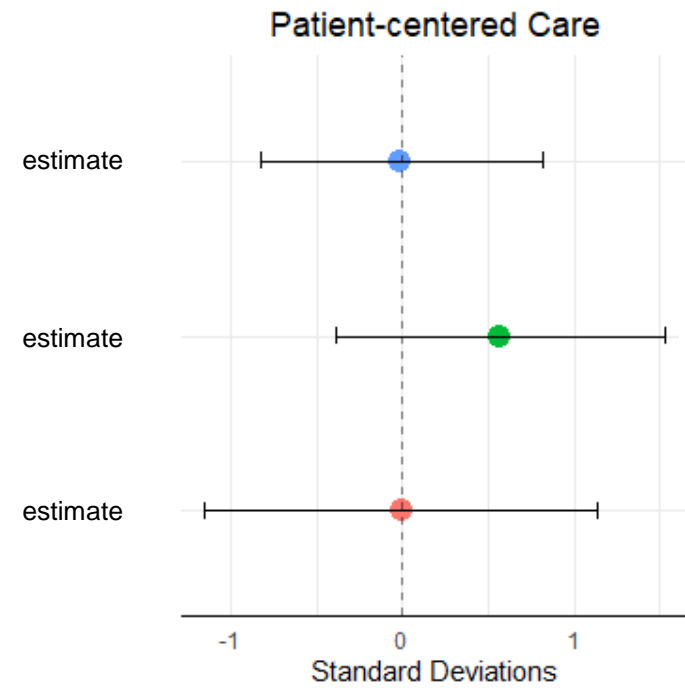
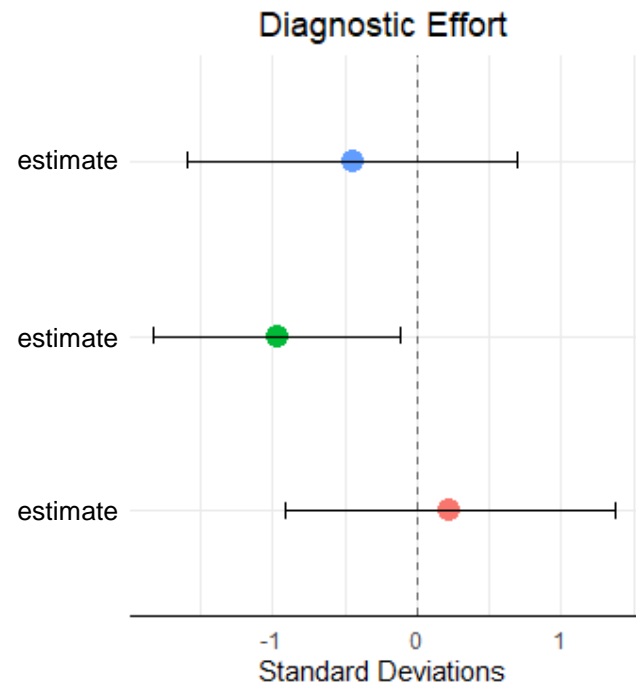
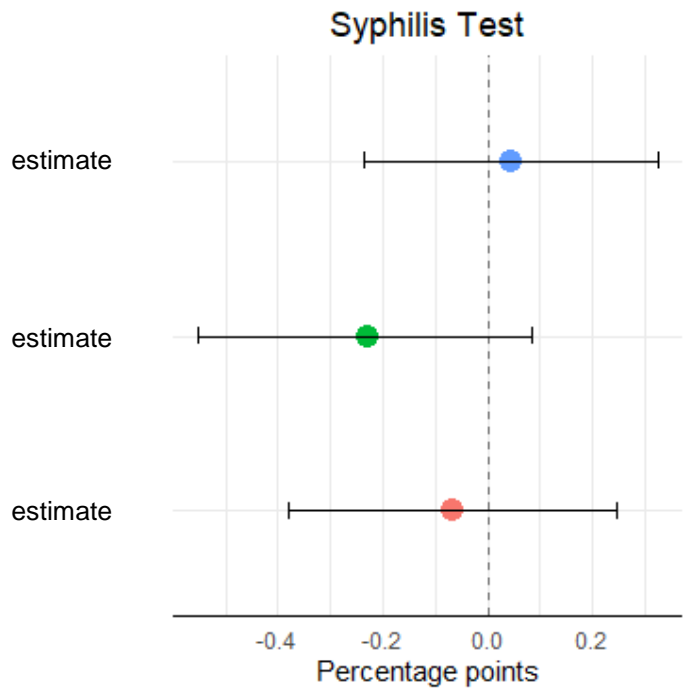
Figure 4. Marginal treatment effects for each of the three primary outcomes. Estimates reflect changes in pre-post stigma measures for each scenario (i.e. HIV negative MSM; HIV positive straight man; HIV positive MSM) relative to the same changes in the referent scenario (HIV negative straight man).

Figure 5. Scenario-specific marginal treatment effects.









- MSM stigma
- HIV stigma
- Intersectional stigma

