

Cluster Randomized Control Trial of the Stepping Stones and Creating Futures Intervention to Prevent Intimate Partner Violence amongst Young People in South Africa: Statistical Analysis Plan

Nada Abdelatif, Esnat Chirwa & Andrew Gibbs

18 September 2018

1. Background

Intimate partner violence (IPV) is a global public health problem. Globally, a third of women are likely to experience IPV in their lifetime and experience of IPV is associated with worse health for women. IPV is exceedingly high in informal settlements in South Africa driven by high levels of poverty, gender inequalities, poor mental health, and substance use. The Stepping Stones and Creating Futures intervention was developed to reduce women's experiences, and men's perpetration of IPV in urban informal settlements in South Africa, through transforming gender norms, and strengthening livelihoods. The study protocol was published in *BMC Public Health* in 2017 and included a short overview of statistical methods. This statistical plan was developed from November 2017 and finalized on 28 September 2018, before database lock.

2. Methods and design

2.1. Study overview

The Stepping Stones and Creating Futures intervention trial is a cluster randomized control trial (CRT), with two arms and is open label. The study population are out-of-school young women and men aged 18-30 (at baseline), resident in informal settlements, and without formal employment. 34 clusters were identified, using both naturally-occurring clusters and researcher-imposed clusters, and then randomized 1:1 by a statistician, blinded to cluster names using numbers.

Because of practical reasons, 16 control clusters had data collected September 2015-December 2015, while 1 control cluster had data collected in August 2016. Phase 1 intervention clusters had data collected January 2016-February 2016, with the intervention delivered until June 2016. Phase 2 intervention clusters had data collected August 2016 – September 2016, with the intervention delivered from then - February 2017. In total 680 women and 677 men were recruited into the study. The study protocol provides further detail on study rationale, randomization, data collection and sample size calculation.

2.2. Intervention

The Stepping Stones and Creating Futures is a participatory intervention, which primarily works with single sex groups. Groups of approximately 20 in size. The intervention seeks to transform gender norms and strengthen livelihoods. It is delivered by same-sex, trained facilitators, and comprises of 21 sessions, each approximately 3 hours long.

2.3. Flow of clusters and participants

The flow of clusters and participants through the trial will be reported using the CONSORT extension for cluster randomized control trials [Figure 1]. The flow diagram will include the number of clusters and cluster randomization, recruitment at baseline by arm, and number of participants at each time point. It will also outline the number of participants withdrawing from the trial at each time point.

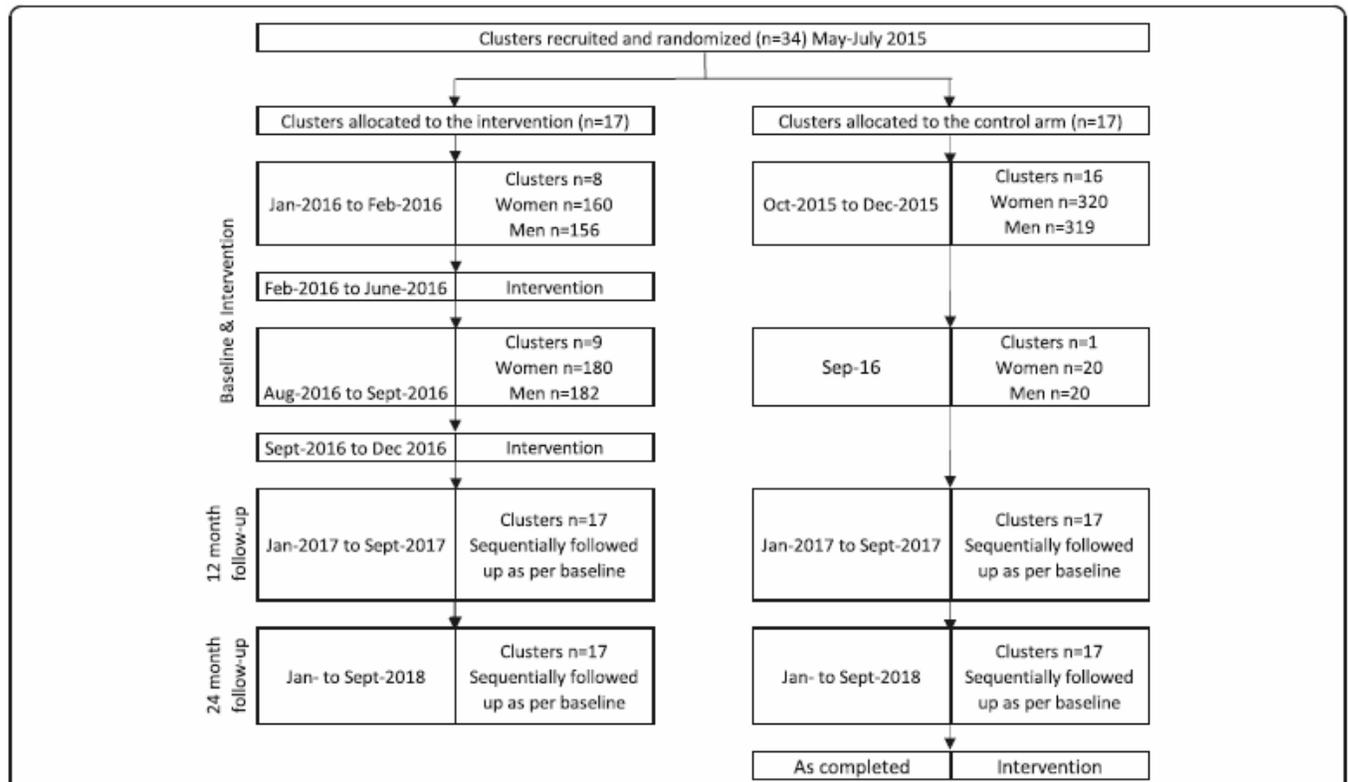


Figure 1: Cluster and participant timeline

2.4. Data collection and management

Randomization was at the cluster level. Once clusters had been recruited, the study statistician randomized clusters using the Microsoft Excel function to randomly allocate equal number of clusters to the intervention and control arm. The statistician was blinded to cluster names. Randomization occurred before recruitment of participants because of political sensitivities.

2.4.1. Data collection

Data was collected by trained fieldworkers. Project Empower (the implementing NGO), in conjunction with external fieldworkers, went to communities and used a community mobilization process to identify potential participants. Over two to three days, through community engagement and snow-ball sampling, the target of 20 women and 20 men were recruited per cluster.

2.4.2. Integrity of data

Data was through self-completed questionnaires on cellphones, in either English, Zulu or Xhosa. The questionnaire had in-built skip patterns and logic checks. Data automatically uploaded to a server and was

downloaded regularly. At 24 months, audio (ACASI) was used to improve understanding by participants. Range and logic checks will be performed on data before analysis.

2.5. Outcomes

2.5.1. Primary outcomes

- Four primary outcomes are linked to IPV:
 - Any past year physical IPV perpetration (men), and experience (women). This is assessed using a modified WHO VAW scale that has been adapted and widely used in South Africa. Five questions are asked about physical IPV perpetration (men) and experience (women) in the past 12 months. Past year physical IPV is coded as positive (1) for anyone responding positively to one or more items on the scale;
 - Any past year sexual IPV perpetration (men), and experience (women) uses the same approach as for physical IPV. Three sexual IPV questions are asked about experiences in the past 12 months. Past year sexual IPV is coded as positive (1) for anyone responding positively to one or more items on the scale;
 - To assess severe past year sexual and/or physical IPV perpetration (men) and experience (women), physical and sexual IPV scales are combined to be a total of eight items. Past year severe sexual and/or physical IPV is assessed as positive if a person responds to two (or more) items once, or one item as few (or more), essentially creating a more than once categorization.
 - Controlling behaviours are assessed using a modified Sexual Relationship Power (SRP) scale [43] with 8 items. Higher scores refer to more controlling behaviours;
- One primary outcome reflects the objective of the Creating Futures component of the intervention, around strengthening livelihoods:
 - Past month earnings are used to assess overall income and livelihoods. A single item question asks, “Considering all the money you earned from jobs or selling things (excluding grants), how much did you earn last month?” Responses are in South African Rands (ZAR) and a continuous scale.

2.5.2. Secondary outcomes

Five groups of secondary outcomes are identified focused on pathways to change through which the intervention is hypothesized to operate.

- The first pathway reflects gender attitudes and norms in the group:
 - Gender attitudes are assessed using a modified Gender Equitable Men’s Scale (GEMS) [44] adapted and widely used in South Africa [18]. The scale comprises of 20 questions, with larger scores indicating less gender equitable attitudes. It is hypothesized the mean will decrease.
- Three sets of questions assess mental health and wellbeing:
 - Past week depressive symptomatology is assessed by the Centre for Epidemiologic Studies Depression (CESD) scale, with the full twenty items [45]. Higher scores indicate greater depressive symptomatology. A mean score for each participant will be calculated. It is hypothesized the mean score will reduce.
 - Past four-week suicidal ideation is assessed using a single item question and a binary yes/no response. It is hypothesised that the percentage of participants reporting yes to this will decrease.

- Life circumstances are assessed using four items derived from the Satisfaction with Life Scale [46]. Higher scores indicate greater life satisfaction. It is hypothesized the mean score will increase.
- Two questions assess the impact of the intervention on alcohol use amongst participants as secondary outcomes:
 - Problem drinking in the past year is assessed using the ten item Alcohol Use Disorders Identification Test (AUDIT) scale. It is hypothesized that there will be a reduction in the mean difference of problem alcohol use.
 - A single item assesses quarrelling in the past year about alcohol consumption with a sexual partner, assessed with a binary yes/no response. The intervention assumes the percentage reporting quarrelling about alcohol use will decrease.
- Two items assess sexual risk behaviour.
 - A single item assesses who the participant last had sex with. Response options are: “main partner, kwapheni (casual) partner, once-off sex partner, ex-partner.” Responses will be coded into a binary of main partner (1) or other sex partner (0). It is hypothesized that the percentage of participants reporting last sex with a main partner will increase.
 - Transactional sex with a kwapheni or once-off sexual partner in the past year will be assessed using a five-item scale used widely in South Africa [48]. A positive response to at least one of these items is classified as responding positively to transactional sex in the past 12 months. It is hypothesized this will reduce.
- Livelihoods are assessed using four scales assessing material outcomes and psychological outcomes of limited income and work opportunities.
 - Shame about lack of work is assessed using four items drawn from the IMAGES study [49]. Higher scores indicate greater levels of shame about lack of work and income. A mean score will be calculated for each participant and it is hypothesized the score will reduce.
 - Stress related to lack of work and income is assessed on a four-item scale drawn from the IMAGES study [49]. Higher scores indicate higher levels of stress about lack of work. A mean score will be calculated for each participant and it is hypothesized the score will reduce.
 - Ability to mobilize cash in an emergency will be assessed with a single item. For analysis a binary will be created through collapsing very difficult and somewhat difficult to indicate challenges (and coded 0), while fairly easy and easy will be coded as no challenge (1). It is hypothesized the percentage of respondents reporting it is fairly easy or easy will increase.
 - Stealing in the past four weeks because of hunger or lack of money will be assessed with a single item. Responses are: Never, once, two or three times, more often. Once, two or three times and more often will be collapsed together for analysis as indicating stealing because of hunger or lack of money. It is hypothesized this will decrease.

This has been summarized in **Table 1**.

Table 1: Primary and secondary trial outcomes

Item	Response categories	Number of items
Primary outcomes		
Physical IPV at 24 months	In the last 12 months how, many times did you push or shove your current or previous girlfriend or wife?	Never, once, few, many 5
Sexual IPV at 24 months	In the last 12 months, how many times have you ever forced your current or previous girlfriend or wife to do something sexual that she did not want to do?	Never, once, few, many 3
Severe IPV at 24 months	Same as above	Never, once, few, many 8
Controlling behaviours at 24 months	I want to know where my partner is all the time.	Strongly disagree, disagree, agree, strongly agree 8
Earnings in the past month at 24 months	Considering all the money you earned from jobs or selling things, how much did you earn in the last 4 weeks (not including grants)?	Continuous variable 1
Secondary outcomes		
Gender attitudes at 24 months	I think that a woman needs her husband's permission to do paid work	Strongly disagree, disagree, agree, strongly agree 20
Depressive symptomology (CESD) at 24 months	During the past week I thought my life had been a failure	Rarely/none of the time, some/little of the time, moderate amount of time, most/all of the time 20
Suicidal ideation at 24 months	In the past four weeks, has the thought of ending your life been in your mind?	Yes, no 1
Life circumstances at 24 months	The conditions of my life are excellent	Strongly disagree, disagree, neither agree or disagree, agree, strongly agree 4
Problem alcohol use (AUDIT) at 24 months	How often in the past year have you had a feeling of guilt or remorse after drinking?	Several items 10
Quarreling about alcohol at 24 months	In the past 12 months have you quarreled with any of your female sexual partners about your drinking?	Yes, no 1
Last sexual partner at 24 months	The last time you had sex was it with a main partner, another partner (khwapeni) or one-off partner or expartner?	Main partner, casual partner, once-off, expartner 1
Transactional sex at 24 months	In the past 12 months, please think about any woman you had sex with just once or any casual partner or khwapeni. Do you think any of them may have become involved	Yes, no 5

	with you because they expected you to give or you gave cash or money to be looked after?		
Work shame at 24 months	I am ashamed to see my girlfriend because I don't have money	Strongly disagree, disagree, agree, strongly agree	4
Work stress at 24 months	I am frequently stressed or depressed because of not having enough work	Strongly disagree, disagree, agree, strongly agree	4
Stealing because of hunger at 24 months	How often in the past 4 weeks have you taken something that was not yours because you did not have enough food or money?	Never, once, two or three times, more	1
Mobilization at cash in emergencies at 24 months	If you had an emergency at home and needed R200, how easy would you say it would be to find the money?	Very difficult, somewhat difficult, fairly easy, easy	1

3. General analysis principles

3.1. Participant population

The primary analysis is an intention to treat (ITT) analysis, based on the assumption that all participants enrolled at baseline will be followed up over the study period and all participants included in the final analysis, and the analysis will be done according to the group they were allocated to, disregarding if they received the intervention.

A small number of participants will be excluded from the analysis. Specifically, if they: 1) were enrolled in school at baseline, 2) if they double-enrolled in the study, 3) if they did not provide primary outcome data at baseline.

3.2. Levels of confidence and p-values

Statistical tests and confidence intervals will be two-sided. The statistical significance level set will be at the 5% level. P-values, confidence intervals and standard errors will be provided.

3.3. Unadjusted and adjusted analysis

All comparative analyses will allow for the clustered nature of the data to ensure correct confidence intervals and type I error rates are calculated. As the trial includes at least 34 clusters, the analyses will be based on the individual-level summarized data. For each outcome, unless otherwise specified, the primary analysis will be the covariate-adjusted analysis, with the statistical models including the stratification variables and baseline values for the outcome under consideration, where available.

Adjustment for baseline covariates is often advised firstly to correct for any chance imbalances in important baseline variables following randomization, and secondly, because adjusting for highly important baseline variables in an RCT can improve the precision of treatment effect estimates even when the outcome measure is binary. Statistical testing for baseline imbalances is not advised and instead key covariates should be selected prior to analysis based on the likely magnitude of the association with the

outcome measure (European Agency for the Evaluation of Medicinal Products, 2003). We have, however, looked at whether there were meaningful differences at baseline by looking at standardized mean differences. We will also perform a multivariable logistic regression to consider any imbalance that may occur in important baseline characteristics known to predict IPV outcomes between the groups using the following variables:

- Baseline term for variable
- Controlling behaviors
- Food insecurity
- Any earnings in past month (none or R1 or more)
- Childhood trauma
- Age
- Education
- Alcohol use

This selection of variables is based on baseline analysis of the data predicting past year physical and/or sexual IPV. Time will also need to be adjusted for since intervention and control clusters were not measured concurrently.

Odds ratios will be quoted together with their 95% confidence intervals and exact P-values. Generalized estimating equations will be used to account for the clustered-nature of the data in the primary analysis (**Section 3.5.1**).

3.4. Missing data

The population is highly mobile and the likelihood of participation retention being less than optimal is likely. To deal with completely missing cases (as opposed to items/variables) a rule has been set in place. If the missing data is $\leq 25\%$ we will use multiple imputation with random-effects logistic regression model approach (Stata v15.1). If there is $>25\%$ missing cases, we will use inverse probability weighting to deal with missing data [refs].

At the individual level, because of data collection techniques, missing data is highly unlikely (in the baseline it was $<1\%$) and as such, we will not impute.

3.5. Proposed Analyses

3.5.1. Primary analysis of primary outcomes

As described above, the primary analysis of the primary outcomes, will follow an intention-to-treat approach, with individuals analyzed according to the trial group to which their settlement was randomized. Since the clusters vary between 17-22 clusters between arms, we will conduct individual-level analysis, as it will be statistically more efficient and allow analysis of the effects of individual-level covariates. Mixed effects models will be used for quantitative outcomes i.e. continuous response variables, to take account of between-cluster variation. For binary outcomes, generalized estimating equations (GEE) will be used with an exchangeable correlation matrix and robust standard errors. GEE performs well for binary data more than the logistic regression random effects model and allow for clustering effects by assuming observations in the same cluster are correlated [Hayes & Moulton, 2017].

3.5.2. Analysis of secondary outcomes

Secondary outcomes will be compared between groups based on the complete data only. Data will be analyzed on an individual-level.

Table 2 provides the proposed analyses for each of the primary and secondary trial outcomes. These analyses will be adjusted for the baseline term of the variable, education and food insecurity.

Table 2: Analysis Methods for trial outcomes

Intention to treat analysis: Primary outcomes			
	Measure	Method of scaling	Proposed analysis
Physical IPV at 24 months	WHO	Binary (never vs once or more)	GEE, odds ratios
Sexual IPV at 24 months	WHO	Binary (never vs once or more)	GEE, odds ratios
Severe IPV at 24 months	WHO	Binary (never or once vs more than once)	GEE, odds ratios
Controlling behaviours at 24 months	Modified SRPS	Mean	Mixed effects linear regression, mean difference
Earnings in the past month at 24 months		Mean	Mixed effects linear regression, mean difference
Intention to treat analysis: Secondary outcomes			
	Measure	Type of variable	Proposed analysis
Gender attitudes at 24 months	Modified GEMS	Mean	Mixed effects linear regression, mean difference
Depressive symptomology at 24 months	CESD	Mean	Mixed effects linear regression, mean difference
Suicidal ideation at 24 months		Binary	GEE, odds ratios
Life circumstances at 24 months		Mean	Mixed effects linear regression, mean difference
Problem alcohol use at 24 months	AUDIT	Mean and binary (cut-off at 7/8)	Mixed effects linear regression, mean difference And GEE, odds ratios
Quarreling about alcohol at 24 months		Binary	GEE, odds ratios
Last sexual partner at 24 months		Binary (main vs other partners)	GEE, odds ratios
Transactional sex at 24 months		Binary (never vs once or more)	GEE, odds ratios
Work shame at 24 months		Mean	Mixed effects linear regression, mean difference
Work stress at 24 months		Mean	Mixed effects linear regression, mean difference

Stealing because of hunger at 24 months		Binary (never stolen vs once or more)	GEE, odds ratios
Mobilization at cash in emergencies at 24 months		Binary (very difficult/somewhat difficult vs fairly easy/easy)	GEE, odds ratios

3.5.3. Additional analyses

To assess the intervention effects at 12 months and 24 months, we will use a mixed effects model at both the cluster and individual level. A sensitivity analysis will be done by performing the analyses with GEE and cluster-level summaries.