

# Affordable Daycare to Empower Indian Women: Pre-Analysis Plan

Sam Harper<sup>1,2</sup>, Anoushaka Chandrashekar<sup>3</sup>, Parul Agarwal<sup>3</sup>, and Arijit Nandi<sup>1,2</sup>

<sup>1</sup>Department of Epidemiology, Biostatistics & Occupational Health, McGill University, Montreal, Canada

<sup>2</sup>Institute for Health and Social Policy, McGill University, Montreal, Canada

<sup>3</sup>Institute for Financial Management Research, Chennai, India

## Abstract

This document outlines the analysis plans for a cluster-randomized evaluation of the impact of a daycare program on women's economic empowerment. We provide an overview of the study background, aims, and primary hypotheses, and we describe in detail the primary and secondary outcomes for women and children. We also describe the primary analyses and models that we will use to evaluate our hypotheses, including an assessment of heterogeneous treatment effects. The primary analyses will be according to intention-to-treat (ITT) but we will also estimate the impact of daycare on women's and children's outcomes using two stage least squares instrumental variable methods.

## 1 Introduction

This document outlines the analysis plans for a cluster-randomized evaluation of the impact of an affordable daycare program on women's economic empowerment. The purpose of a pre-analysis plan is to provide a clear and transparent accounting of measurements, hypotheses, and statistical analyses that will be conducted prior to the investigators having access to data on outcomes. The registration and publication of pre-analysis plans is hoped to reduce problems related to data mining, publication bias, and missing trial data,<sup>1</sup> in addition to prompting researchers to think critically about measurement and analytic issues prior to any actual data analysis.<sup>2</sup> Nevertheless, such plans are challenging to fully specify in advance, especially for complex interventions that may have impacts on multiple outcomes.<sup>3</sup> This pre-analysis plan was drafted after the collection of baseline data, but while data collection from the first follow-up survey was ongoing and data were inaccessible to the authors. Subsequent data analysis will follow the plan laid out in this manuscript and any deviations from the plan will make reference to this document for transparency.

### 1.1 Background

The *Uttam Unnati* or "great progress" study is a cluster-randomized evaluation of the impact of an affordable daycare program on the well-being of mothers and children taking place in Rajasthan, India. The protocol, detailed design and sampling information, and power calculations for this study have been published previously.<sup>4</sup> The trial is registered in the International Standard Randomised Controlled Trials Number clinical trial registry (ISRCTN45369145, <http://www.isrctn.com/ISRCTN45369145>) and in the American Economic Association's registry for

randomized controlled trials (AEARCTR-0000774, <http://www.socialscienceregistry.org/trials/774>). In this manuscript we provide additional details on the primary hypotheses, outcome measures, and pre-specified analyses that will be conducted after the first round of follow-up surveys have been completed.

## 1.2 Aims

The primary aims of the study are to provide evidence on whether providing access to an affordable, community-based day care program affects women's social and economic well-being, including their health. Additionally, we aim to provide evidence of program impacts on children's health and schooling. The primary study hypotheses are that providing access to affordable daycare will: 1) improve measures of women's social and economic empowerment; 2) reduce the time women spend directly caring for children; 3) improve women's mental health; and 4) improve the health, nutrition, and schooling of children. Below we provide additional details on the intervention and the specific outcome measures we will use to test these hypotheses.

## 2 Experimental Design

This study takes place in the Udaipur district of Rajasthan, India. For this study we used a cluster-randomized design with stratification by geographic area, specifically 5 blocks within the Udaipur District. Because daycare programs are made available to the entire community it was not feasible to randomize at the level of individual families because of potential contamination. Instead, we randomized at the level of the village hamlet (hamlets are small settlements, formed by community members, that surround a village center but constitute independent units). Our sample included 160 village hamlets. Based on focus group data and prior knowledge of heterogeneity in the Udaipur district, we chose to stratify by block group to prevent variations in the distributions of blocks across treatment groups (e.g., if women in treated hamlets were more likely to reside in blocks with more economic opportunities). Because four of the five blocks contained an odd number of hamlets, the randomization was done so that two of the four blocks would have an additional treated hamlet and the remaining two would have an additional control hamlet.

Due to the nature of the intervention it was not possible to blind study personnel or participants after the implementation of the intervention. However, we concealed the allocation of hamlets to treatment or control status until after the baseline survey in order to minimize opportunities for bias in recruitment of participants and the baseline survey.

### 2.1 Intervention

The intervention was implemented by a local non-governmental organization called Seva Mandir. Briefly, in each of 80 treated clusters, a program of full time, affordable, community-run daycare was established. The daycare centers are called balwadis, and provide full day child-care, nutritious food and supplements, basic medicines, and preschool education to children one to six years old. Additionally, the balwadis support increasing child immunization coverage by maintaining immunization records and creating awareness among parents and connecting them with government nurses when required. The centers are operated by trained local women, called sanchalikas, who are hired and trained by Seva Mandir. The sanchalikas meet quarterly with parents to provide feedback on child progress. To encourage community enrollment in the program, the balwadi program was also accompanied by a household marketing campaign.

## 2.2 Monitoring

Independent fieldworkers employed by the Institute for Financial Management and Research are conducting monthly site visits to each balwadi to verify that the balwadi is operating and the structure is adequate, and that the sanchalika is providing sufficient food and education to the children. Data from the monthly site visits will be used to supplement information gathered from a camera-monitoring system, which will be used to measure the number of days each balwadi is operating and any reasons for temporary closings. Sanchalikas are required to take three self-timed photos each day, which are used to evaluate the number of days when the balwadi was run for at least 6 hours, which is considered a full day of operation. We will use the monitoring data to assess the impact of treatment assignment on actual receipt and use of daycare among women and children.

## 3 Survey methods

In late 2014 we completed a household census in village hamlets in the Udaipur district in Rajasthan, India to confirm the eligibility of the hamlet, enumerate the population, and identify potential respondents for inclusion. We conducted a baseline survey in early 2015 among 3177 mothers with a child between one and six years of age living in 160 village hamlets from five blocks. After describing the study objectives, procedures, potential risks, potential benefits, voluntary nature, confidentiality and privacy protections, and compensation, each eligible respondent was asked if they would consent to participate. This was done in written form for respondents who could read and write and orally for those who could not. Each respondent who agreed to participate received a blanket, valued at 100 rupees (Rs.), as compensation for their time at the completion of the survey interview.

The survey questionnaire was translated from English to the local Hindi language by a professional translator, and the software for recording responses via netbook was developed by the Institute for Financial Management and Research—Leveraging Evidence for Access and Development (IFMR-LEAD) software team. To minimize error a survey supervisor observed each surveyor for one survey each day, monitoring for correctness in asking survey questions, providing necessary clarifications, and responding to respondent questions. A random set of 10 % of completed surveys was selected for back-checks—a list of questions asked in the survey that are unlikely to change over a short time period, like the number of children, respondent's marital and educational status and employment history. The back-check team independently re-interviewed the respondents and entered their responses into a software program designed to test these questions for identical responses. A third party clarified any discrepancies between the original survey response and the back-check response. More detail on specific outcome measures and covariates are provided below.

## 4 Measures

As noted above, our primary interest is in whether the intervention affects measures of women's economic and social empowerment, and on the health and nutrition of their children. Here we provide a summary of the rationale for the choice of primary and secondary study outcomes for mothers and children.

## 4.1 Outcomes for mothers

**Primary outcomes** The primary outcomes for women include women’s employment, income and savings, and measures of empowerment. Based on prior literature suggesting that access to daycare increases women’s labor force attachment<sup>5-9</sup> and income,<sup>10,11</sup> we hypothesize that women living in treated hamlets would be more likely to engage in non-domestic work after the intervention. We collected information on the type of work that women were performing (examples?), and will investigate the program impact for specific categories of work. Opportunities for paid employment in Rajasthan are limited, as indicated by the small fraction (8%) of women who reported that they worked for cash in the past week at baseline, so we anticipate a modest increase in paid employment.

To the extent that child care relief may provide women with additional opportunities for work, we will also examine whether the treatment affects women’s income and savings. We collected detailed information on the household income received in the past 12 months from various categories (e.g., agricultural income, business income, rents, remittances, government payments), in addition to savings accounts held by household members, including for each account the type of account, its purpose, the total value, and whether the respondent can use the account to make purchases. We will also assess whether the intervention reduces the probability of a household reporting that they have a Below Poverty Line (BPL) card, India’s scheme that identifies households in need of government assistance and aid.

India’s National Rural Employment Guarantee Act (NREGA) provides at least 100 days of wage employment to rural adults that are willing to do unskilled manual work and provides work opportunities to our participants. Access to daycare should provide increased opportunities to take advantage of NREGA so we will also assess whether the treatment influenced the probability of NREGA participation. The treatment could affect household income and savings independently of women’s formal labor force participation if access to daycare influences husband’s earning opportunities. Table 1 shows the primary outcomes and how they will be measured.

**Secondary outcomes** Secondary study outcomes for women include their use of time, self-rated mental health, and experience of intimate partner violence. Time use is one of the mechanisms by which access to daycare might influence the primary outcomes of interest, including women’s economic empowerment. We asked respondents whether they spent any time in the past 24 hours on specific activities (e.g., gathering firewood, laundry), how much time they spent on each activity, and whether that was the usual amount of time for that activity. We will assess the intervention impact on specific categories of time allocation: paid work, unpaid household work (e.g., caring for children), and unpaid non-household work (e.g., farming).

By reducing conflicting demands on women’s time, we also expect that the treatment has the potential to reduce women’s levels of mental distress. Focus group discussions prior to the study suggested that mothers felt particular distress related to the safety of their children while conducting routine household activities. We used the Hindi version of the 12-item General Health Questionnaire (GHQ-12), translated by Gautam et al.,<sup>12</sup> to measure symptoms of mental distress (e.g., “Have you recently been able to concentrate on what you’re doing?”). Responses for each of the 12 questions range from 1 to 4 on a Likert scale.

We also will assess the program’s impact on intimate partner violence. There are reasons to expect that daycare might have either positive or negative effects on intimate partner violence. If access to daycare improves household dynamics and interpersonal relationships it could reduce a husband’s controlling behaviors and perpetration of physical and emotional abuse. A few studies, however, have examined the impact of women’s empowerment programs on intimate partner violence;<sup>13</sup> there are mixed results, with some literature indicating that these programs might actually exacerbate

domestic violence.<sup>14</sup> If daycare were to empower women to assert themselves in decision-making processes that threaten the husband's authority, then it could result in the increased incidence of intimate partner violence. Measures of secondary outcomes for women are provided in Table 2.

## 4.2 Outcomes for children

**Primary outcomes** The main outcomes for children included their nutritional status, the incidence of specific illnesses and use of healthcare, and vaccination status. We anticipated that the treatment, by providing access to a daycare program that includes healthy meals, would improve children's nutritional status, including shorter and longer-term indicators. The treatment could influence the incidence of illness in either direction. If the treatment induced mothers to join the labor force, for example, then they might be less available to take their children to see a health professional in the case of illness. Conversely, if access to daycare improved household economic conditions then families may be able to afford out-of-pocket health expenditures, which may have been a barrier to accessing care. The program is also intended to track children's vaccination status and refer children for vaccination services so we hypothesized that the treatment would increase immunization coverage. Primary outcomes and measures for children are listed in greater detail in Table 3.

**Secondary outcomes** Secondary outcomes among children include schooling and literacy. For all children in the household (not just those who enroll in daycare) we will track their school attendance, including the average number of days attended per week and average number of hours per day on the days attended. Young children (especially young girls) often spend time helping to care for their younger siblings, and this may be one reason why young girls are more likely than boys to drop out of school. We hypothesize that daycare provision may help relieve the burden on younger siblings and could lead to increased attendance and duration of schooling. Because the balwadi program also includes a consistent educational curriculum, we also hypothesize that children in treated areas may achieve higher levels of literacy and numeracy, which will be measured by mother's self-reports of whether the child can read or write.

## 5 Statistical Analysis

Our basic framework is a cluster-randomized evaluation, with clusters specified at the level of the hamlet. A key issue for inference from such a design is to account for non-independence among observations from the same cluster when estimating the impact of the intervention.<sup>15</sup> We focus here on the general plan for analysis of survey results, including descriptive tables and basic model specifications.

### 5.1 Descriptive tables

We will generate a descriptive table of measures of central tendency and dispersion for the main outcome variables and the primary covariates specified above, and provide both cluster-level summaries and individual-level summaries.<sup>16</sup> Continuous variables will be presented as means and standard deviations, and categorical variables as proportions and standard deviations. We will also provide a breakdown by our main stratification variable, block. We will refrain from conducting statistical tests of balance across treated and control groups.<sup>17</sup>

## 5.2 Sample attrition and missing data

Given the relatively high initial response rate for our baseline survey (~90%), relatively short follow-up, and generally limited long-term migration in the study area, we anticipate little study attrition. Even for those observed at follow-up it is also possible that some primary or secondary outcome data could be missing, although we anticipate relatively little item non-response, based on the baseline survey results. Regardless, sample attrition and missing data are common in virtually all surveys and we describe briefly our plan for dealing with these issues. If the proportion of missing outcome data is relatively small (e.g., <5%) we will proceed with a complete case analysis. Regardless of the amount of missing data, we will describe the reason why subjects were lost to follow-up and why outcomes are missing.<sup>18,19</sup> In addition to reporting baseline characteristics we will also include characteristics of those who were lost to follow-up or were missing outcome or covariate data. For amounts of missing data that may be more severe we will perform sensitivity analysis to quantify the effect of missing outcome data on our effects of interest (using multiple imputation, inverse probability weighting, or estimating bounds for our treatment effects).<sup>20-22</sup>

## 5.3 Main model specifications

The primary specification will be an intent-to-treat (ITT) model including adjustment for stratification by block. For continuous outcomes we will use linear model of the form:

$$y_{ij} = \alpha + \beta Z_j + \sum_{k=1}^K \gamma_k K_k + \varepsilon_{ij} \quad (1)$$

where  $y_{ij}$  is the outcome of interest for individual  $i$  in cluster  $j$ ,  $Z_j$  is the main cluster-level treatment assignment variable (1 for intervention, 0 for control),  $\gamma_k$  are coefficients for each of the  $K$  blocks on which we stratified randomized treatment assignment, and  $\varepsilon_{ij}$  is an error term. To account for non-independence of errors among observations in the same hamlet we will use cluster-robust variance estimators in all models.<sup>15,23</sup> The coefficient  $\beta$  and its 95% confidence interval will be our primary estimate of interest. Equation 1 above represents our unadjusted model, but in order to increase precision we may include additional baseline covariates that are strong predictors of the outcome. It is often the case that one of the strongest predictors of post-treatment outcome is pre-treatment outcome measured at baseline, and we will include this as well:

$$y_{ij} = \alpha + \beta Z_j + \sum_{k=1}^K \gamma_k K_k + \sum_{m=1}^M \delta_m C_{ijm} + \varepsilon_{ij} \quad (2)$$

where now  $C$  is a vector of  $M$  pre-treatment individual-level covariates, including a measure of the outcome  $y_{ij}$  at baseline. In particular, including measures of the outcome at baseline can lead to meaningful increases in power, especially for outcomes that are weakly correlated over time.<sup>24,25</sup> For continuous outcomes (e.g., empowerment scale) we will use linear regression. For binary outcomes (e.g., maternal employment) we will use logistic regression and report marginal effects on the absolute risk scale. For count outcomes (e.g., reports of symptoms of distress) we will use either Poisson or negative binomial regression and report marginal estimates of the expected number of events. For all models we will use cluster-robust standard errors, clustered at the hamlet level.<sup>23</sup>

## 5.4 Heterogeneous treatment effects

Prior literature has found some evidence that access to center-based child care had stronger impacts on mothers who may have been more disadvantaged.<sup>7,9</sup> Therefore, we will assess whether there are any differences in the program impact according to the following maternal characteristics: 1) Current work status (whether or not women were working at baseline); 2) Education (less than primary vs. primary or greater); 3) Household structure (e.g., whether there is another adult caregiver in house); 4) Measures of women’s empowerment at baseline (continuous latent scale from confirmatory factor analysis); and 5) Distance from the village center. Additionally, we will test whether the program has any differential impact by our main geographic stratification variable (block). For outcomes among children we will also test whether the program has differential effects by baseline measures of child health. In order to test whether the program effects may differ by other individual characteristics, we will extend the generalized linear model specified with covariates (including block) above (Equation 2) to allow for heterogeneity in the impact of the treatment:

$$y_{ij} = \alpha + \beta Z_j + \sum_{k=1}^K \gamma_k K_k + \sum_{m=1}^M \delta_m \mathbf{C}_{ijm} + \theta (Z_j \times C_{ij}) + \varepsilon_{ij} \quad (3)$$

where the coefficient  $\theta$  on the product term between treatment assignment ( $Z_j$ ) and a specific covariate  $C$  provides a test of whether the program’s impact is homogenous across levels of  $C$ . The other terms in Equation 3 are defined analogously to Equation 2, though the inclusion of the product term affects their interpretation. Based on prior research we anticipate stronger treatment effects for women who were not working at baseline, with low education, without another adult caregiver in the household, and with higher baseline levels of empowerment.

## 5.5 Multiple comparisons

Our primary analytic goal for this study is to estimate treatment effects and measures of precision (i.e., confidence intervals) for the intervention, and we will generally avoid null hypothesis significance testing (with the exception of tests for heterogeneity as described above). However, because daycare is a complex intervention that may plausibly affect a number of outcomes for both mothers and children, in order maintain an overall Type 1 error rate we will make some adjustments for testing the impact of the intervention across many outcomes.<sup>26–28</sup> To test whether the intervention has general impacts on a given domain with multiple indicators (e.g., empowerment, employment, income/savings, time use) we will create a summary index for each domain. For related outcomes within a given domain we will take the approach of creating standardized effect estimates for each of the subdomains, averaging them, and calculating a summary estimate that accounts for testing multiple outcomes.<sup>29,30</sup> The use of an index is beneficial in this case since probability of a false rejection does not increase as additional outcomes are added to a summary index. Moreover, an indexed outcome can provide a statistical test for whether a program has a “general effect” on a set of outcomes.

Using an index, however, will not provide us with inference on specific sub-components within a given domain (e.g., a women’s decision-making ability regarding her own health care in the empowerment domain). Within a given domain we will report both unadjusted  $p$ -values for individual outcomes and “family-wise”  $p$ -values adjusted to account for the multiple outcomes examined within a domain.<sup>30,31</sup> To avoid the losses of power associated with simple Bonferroni-type adjustments we will use the free step-down method of Westfall and Young.<sup>32</sup> Finally, because many of our tests for heterogeneous effects are exploratory, we will use the less conservative strategy of Benjamini and Hochberg<sup>33</sup> account for multiple testing across the standardized domain outcomes.

## 5.6 Estimating the impact of use of daycare on maternal outcomes

All of the prior analyses will estimate the impact of random assignment to daycare. However, we are also interested in estimating the impact of actual receipt of daycare on maternal and child outcomes. For these analyses the basic identifying assumption is that random assignment to daycare only affects outcomes through its impact on the actual take-up and use of daycare. Equation 1 above may be considered as the ITT or reduced form equation for the impact of treatment assignment, and we will use two-stage least squares (2SLS) instrumental variables analysis to estimate the impact of daycare on maternal outcomes. We will proceed by estimating the first stage equation, which gives the impact of randomized treatment assignment on daycare:

$$D_{ij} = \alpha + \beta Z_j + \sum_{k=1}^K \gamma_k K_{jk} + \sum_{m=1}^M \delta_m C_{ijm} + \varepsilon_{ij} \quad (4)$$

where  $D_{ij}$  now represents whether or not mother reports enrolling her child in the balwadi program and the other parameters are defined similarly as in Equation 1 above. We use the coefficients from Equation 3 to predict, for each observation, the probability  $\hat{D}_{ij}$  of enrolling in the balwadi program, which is then used in the second stage equation:

$$y_{ij} = \zeta + \eta \hat{D}_{ij} + \sum_{k=1}^K \theta_k K_{jk} + \sum_{m=1}^M \lambda_m C_{ijm} + \varepsilon_{ij} \quad (5)$$

where  $y_{ij}$  is the outcome of interest,  $\hat{D}_{ij}$  is the predicted probability of daycare use derived from Equation 4 and the coefficient. We will use software to correct standard errors for the fact that  $\hat{D}_{ij}$  is predicted in the first stage.<sup>34,35</sup>

## 6 Conclusion

Access to affordable daycare may have important impacts on the lives of the women and children of rural India. This pre-analysis plan makes clear what our main hypotheses are regarding the intervention's impact on women and children, how we will measure primary and secondary outcomes, and how we will analyze the evaluation data.

## References

- [1] Ioannidis JPA. How to make more published research true. *PLoS Med.* 2014 Oct;11(10):e1001747.
- [2] Miguel E, Camerer C, Casey K, Cohen J, Esterling KM, Gerber A, et al. Social science. Promoting transparency in social science research. *Science.* 2014 Jan;343(6166):30–1.
- [3] Olken BA. Promises and perils of pre-analysis plans. *The Journal of Economic Perspectives.* 2015;29(3):61–80.
- [4] Nandi A, Maloney S, Agarwal P, Chandrashekar A, Harper S. The effect of an affordable daycare program on health and economic well-being in Rajasthan, India: protocol for a cluster-randomized impact evaluation study. *BMC Public Health.* 2016 Jun;16:490.

- [5] Attanasio O, Vera-Hernandez M. Medium- and long run effects of nutrition and child care: evaluation of a community nursery programme in rural Colombia, IFS report EWP04/06. The Institute for Fiscal Studies; 2004.
- [6] Berlinski S, Galiani S, Gertler P. The effect of pre-primary education on primary school performance [Journal Article]. *Journal of public Economics*. 2009;93(1):219–234.
- [7] Barros R, Olinto P, Lunde T, Carvalho M. The impact of access to free childcare on women’s labor market outcomes: evidence from a randomized trial in low-income neighborhoods of Rio de Janeiro. In: *World Bank Economists’ Forum*; 2011. .
- [8] Berlinski S, Galiani S, McEwan PJ. Preschool and maternal labor market outcomes: evidence from a regression discontinuity design [Journal Article]. *Economic Development and Cultural Change*. 2011;59(2):313–44.
- [9] Angeles G, Gadsden P, Galiani S, Gertler P, Herrera A, Kariger P, et al. The impact of daycare on maternal labour supply and child development in Mexico, 3ie Impact Evaluation Report 6. International Initiative for Impact Evaluation (3ie); 2014.
- [10] Deutsch R. Does child care pay. Labor-force participation and earnings effects of access to childcare in the Favelas of Rio de Janeiro The Inter-American Development Bank, Washington, DC Photocopy. 1998;
- [11] Nakahara S, Poudel KC, Lopchan M, Poudel OR, Poudel-Tandukar K, Ichikawa M. Differential effects of out-of-home day care in improving child nutrition and augmenting maternal income among those with and without childcare support: A prospective before–after comparison study in Pokhara, Nepal. *Health policy*. 2010;97(1):16–25.
- [12] Gautam S, Nijhawan M, Kamal P. Standardisation of hindi version of Goldberg’s General Health Questionnaire [Journal Article]. *Indian journal of psychiatry*. 1987;29(1):63.
- [13] Vyas S, Watts C. How does economic empowerment affect women’s risk of intimate partner violence in low and middle income countries? A systematic review of published evidence. *Journal of international Development*. 2009;21(5):577–602.
- [14] Weitzman A. Women’s and Men’s Relative Status and Intimate Partner Violence in India. *Population and Development Review*. 2014;40(1):55–75.
- [15] Hayes RJ, Moulton LH. Cluster randomised trials. *Interdisciplinary statistics series*. Boca Raton: CRC Press; 2009. Available from: <http://www.loc.gov/catdir/toc/ecip0826/2008035804.html>.
- [16] Campbell MK, Elbourne DR, Altman DG. CONSORT statement: extension to cluster randomised trials [Journal Article]. *Bmj*. 2004;328(7441):702–708.
- [17] Senn S. Testing for baseline balance in clinical trials. *Stat Med*. 1994 Sep;13(17):1715–26.
- [18] The prevention and treatment of missing data in clinical trials. Washington, D.C.: National Academies Press; 2010. Available from: <http://www.nap.edu/catalog.php?record%5Fid=12955>.
- [19] Groenwold RHH, Moons KGM, Vandenbroucke JP. Randomized trials with missing outcome data: how to analyze and what to report. *CMAJ*. 2014 Oct;186(15):1153–7.
- [20] Manski CF. Nonparametric bounds on treatment effects. *The American Economic Review*. 1990;80(2):319–323.
- [21] Lee DS. Training, wages, and sample selection: Estimating sharp bounds on treatment effects. *The Review of Economic Studies*. 2009;76(3):1071–1102.

- [22] White IR, Horton NJ, Carpenter J, Pocock SJ. Strategy for intention to treat analysis in randomised trials with missing outcome data. *BMJ*. 2011;342:d40.
- [23] Cameron AC, Miller DL. A practitioner's guide to cluster-robust inference. *Journal of Human Resources*. 2015;50(2):317–372.
- [24] Frison L, Pocock SJ. Repeated measures in clinical trials: analysis using mean summary statistics and its implications for design. *Stat Med*. 1992 Sep;11(13):1685–704.
- [25] McKenzie D. Beyond baseline and follow-up: The case for more T in experiments. *Journal of Development Economics*. 2012;99(2):210–221.
- [26] Schochet PZ. Technical Methods Report: Guidelines for Multiple Testing in Impact Evaluations. NCEE 2008-4018. National Center for Education Evaluation and Regional Assistance. 2008;
- [27] Anderson ML. Multiple inference and gender differences in the effects of early intervention: A reevaluation of the Abecedarian, Perry Preschool, and Early Training Projects. *Journal of the American statistical Association*. 2012;
- [28] Glennerster R, Takavarasha K. *Running randomized evaluations: a practical guide*. Princeton, NJ: Princeton University Press; 2013.
- [29] O'Brien PC. Procedures for comparing samples with multiple endpoints. *Biometrics*. 1984 Dec;40(4):1079–87.
- [30] Kling JR, Liebman JB, Katz LF. Experimental analysis of neighborhood effects. *Econometrica*. 2007;75(1):83–119.
- [31] Finkelstein A, Taubman S, Wright B, Bernstein M, Gruber J, Newhouse JP, et al. The Oregon Health Insurance Experiment: Evidence from the First Year. *The Quarterly journal of economics*. 2012;127(3):1057–1106.
- [32] Westfall PH, Young SS. *Resampling-based multiple testing: examples and methods for P-value adjustment*. New York: Wiley; 1993. Available from: <http://www.loc.gov/catdir/description/wiley031/92022194.html>.
- [33] Benjamini Y, Hochberg Y. Controlling the false discovery rate: a practical and powerful approach to multiple testing. *Journal of the royal statistical society Series B (Methodological)*. 1995;p. 289–300.
- [34] Angrist JD, Imbens GW, Rubin DB. Identification of causal effects using instrumental variables. *Journal of the American Statistical Association*. 1996;91(434):444–455.
- [35] Angrist JD, Pischke JS. *Mostly harmless econometrics: an empiricist's companion*. Princeton Univ Press; 2008.

**Table 1:** Primary outcomes and measures for women

Primary outcome category	Dimension	Measure
Empowerment	Household decision making	Constructed from 8 questions pertaining to women's involvement in household decisions (i.e., decisions about: her own healthcare, whether she can work, where she can work, education of her children, visits to family and friends, making major household purchases, making minor household purchases, and how husband's earnings are spent). Potential answers are: mainly you, mainly your husband, you and your husband jointly, or someone else. We will sum up the number of decisions women were involved in (1 if woman has any involvement, potential score range: 0-8).
	Freedom of movement	Summary score of binary responses to places within the village women could travel alone (place of worship, friends' homes, market, health center, community center). Potential responses: alone, not alone, not at all. Scored 1 if women may travel along (potential score range: 0-5).
	Community participation	Regular attendance at group
	Domestic violence attitudes	Agrees husband is justified in hitting his wife in any the following situations: 1) She leaves the house without telling him; 2) She neglects the house or the children; 3) She argues with him; 4) She doesn't cook food properly; or 5) She shows disrespect for her in-laws.
	Gender norm attitudes	Endorses either: 1) a married woman should be allowed to work outside the home if she wants to; or 2) It is better to send a son to school than a daughter.
	Continuous index	Scaled score based on confirmatory factor analysis.
Employment	Any work	Did you do any work in the past month? (yes/no)
	Paid work	Are you paid in cash or kind for this work (yes/no)
	Working days	Number of days worked in a typical week for a wage
	NREGA participation	Paid through NREGA (yes / no)
	Daycare utilization	Leaves child at formal daycare when working (yes / no)
Income and Savings	Own income	Income reported from all sources (rupees)
	Control over own savings	Woman can make purchases without husband's permission (yes / no)
	Poverty status	Has below poverty line (BPL) card.
	Household wealth	Principle component analysis (PCA) using 27 indicators including housing characteristics, the number of durables owned, property ownership, and whether the household had a savings account.

**Table 2:** Secondary outcomes and measures for women

Secondary outcome category	Dimension	Measure
Mental distress	Count of symptoms	Number of symptoms of distress from GHQ-12
Time use	Paid work	Minutes spent on paid agricultural labor, non-agricultural labor, other paid work (past 24 hours)
	Unpaid household work	Minutes spent collecting water, cooking, cleaning, laundry, caring for others, and gathering fuel (past 24 hours)
	Unpaid non-household work	Minutes spent tending animals or working in own field (past 24 hours)
Intimate partner violence	Physical abuse	Positive response to any of: slapped; twisted arm/pulled hair; pushed, shook, threw something; punched; kicked or dragged; choked; threatened or attacked with weapon in past 12 months.
	Emotional abuse	Positive response to any of: 1) Husband says something to humiliate you in front of others; 2) Husband threatens to hurt or harm you or someone close to you; 3) Husband insults you or makes you feel bad about yourself in last 12 months.
	Controlling behavior	Positive response to any one of 5 controlling behaviors (jealous when talking to other men, limiting contact with female friends, limiting contact with family, insistence on knowing location, distrust with money).
	Any type of violence	Yes to any of the above sub-categories of violence.

**Table 3:** Primary outcomes and measures for children

Primary outcome category	Dimension	Measure
Daycare utilization	Balwadi attendance	Attended balwadi in past year (yes/no).
	Balwadi frequency	Days per week attending balwadi.
	Balwadi duration	Hours per day attending balwadi.
Nutrition	Underweight	Below minus two standard deviations from median weight-for-age of WHO reference population.
	Stunting	Below minus two standard deviations from median height-for-age of WHO reference population.
	Wasting	Below minus two standard deviations from median weight-for-height of WHO reference population.
Incidence of illness symptoms	Fever or chills	Mother's report of child symptoms in past month.
	Persistent cough	Mother's report of child symptoms in past month.
	Diarrhea	Mother's report of child symptoms in past month.
	Injury	Mother's report of child symptoms (broken bone, cut with heavy bleeding, skin burn) in past month.
	Any illness	Mother's report of any of the above categories in past month.
Vaccination	Tuberculosis	Child received BCG (tuberculosis) vaccine.
	Polio	Child received 3 doses of polio vaccine.
	Diphtheria and tetanus	Child received 3 doses of diphtheria and tetanus vaccine.
	Measles	Child received measles vaccine.
	Fully vaccinated	Child received BCG (tuberculosis) and Polio (3 doses) and diphtheria and tetanus (3 doses) and measles vaccines.

**Table 4:** Secondary outcomes and measures for children

<b>Secondary outcome category</b>	<b>Dimension</b>	<b>Measure</b>
Schooling	School attendance	Whether child currently attending school this year (yes/no).
	School frequency	Number of days per week attending school.
	School duration	Number of hours per day attending school.
	Older sibling schooling	Whether older sibling attending school (yes/no).
Literacy	Reading	Maternal report (yes / no) of whether the child can read.
	Writing	Maternal report (yes / no) of whether the child can write.