Impact evaluation of the Care Tipping Point Initiative in Nepal: study protocol for a mixed-methods cluster randomised controlled trial

Kathryn M Yount, Cari Jo Clark, Irina Bergenfeld, Zara Khan, Yuk Fai Cheong, Sadhvi Kalra, Sudhindra Sharma, Shuvechha Ghimire, Ruchira T Naved, Kausar Parvin, Mahfuz Al Mamun, Aloka Talukder, Anne Laterra, Anne Sprinkel, on behalf of the Tipping Point Program Study Team

ABSTRACT

Introduction Girl child, early and forced marriage (CEFM) persists in South Asia, with long-term consequences for girls. CARE’s Tipping Point Initiative (TPI) addresses the causes of CEFM by challenging repressive gender norms and inequalities. The TPI engages different participant groups on programmatic topics and supports community dialogue to build girls’ agency, shift inequitable power relations, and change community norms sustaining CEFM.

Methods/analysis The Nepal TPI impact evaluation has an integrated, mixed-methods design. The quantitative evaluation is a three-arm, cluster randomised controlled trial (control; Tipping Point Programme (TPP); TPP+ with emphasised social norms change). Fifty-four clusters of ~200 households were selected from two districts (27:27) with probability proportional to size and randomised. A household census ascertained eligible study participants, including unmarried girls and boys 12–16 years (1242:1242) and women and men 25+ years (270:270). Baseline participation was 1134 girls, 1154 boys, 270 women and 270 men. Questionnaires covered agency; social networks/norms; and discrimination/violence. Thirty-in-depth interviews, 8 key-informant interviews and 32 focus group discussions were held across eight TPP/TPP+ clusters. Guides covered gender roles/aspirations; marriage decisions; girls’ safety/mobility; collective action; perceived shifts in child marriage; and norms about girls. Monitoring involves qualitative interviews, focus groups and session/event observations over two visits. Qualitative analyses follow a modified grounded theory approach. Quantitative analyses apply intention to treat, regression-based difference-in-difference strategies to assess impacts on primary (married, marriage hazard) and secondary outcomes, targeted endline tracing and regression-based methods to address potential selection bias.

Ethics/dissemination The Nepal Social Welfare Council approved CARE Nepal to operate in the study districts. Emory (IRB00109419) and the Nepal Health Research Council (161–2019) approved the study. We follow UNICEF and CARE guidelines for ethical research involving children and gender-based violence. Study materials are here or available on request. We will share findings through clinicaltrials.gov, CARE reports/briefs and publications.

Strengths and limitations of this study

- The Tipping Point Initiative addresses the causes of child, early and forced marriage (CEFM).
- The Nepal Tipping Point Impact Evaluation has an integrated, mixed-methods design.
- The qualitative longitudinal study provides insights about the normative context of CEFM.
- The three-arm cluster-randomised controlled trial assesses incremental effects of norm-change programming.
- Given challenges with programme recruitment, demographics will be compared with local census data and inferences made to the study sample.

INTRODUCTION

Child, early and forced marriage: prevalence and causes

Each year, child, early and forced marriage (CEFM), usually defined as marriage before age 18 years, affects more than 10 million girls globally.1 These marriages violate girls’ rights to bodily autonomy, health, education and opportunity. CEFM also results in lifelong consequences for the physical, emotional, material and psychological well-being of girls.2–4 About half of all child marriages occur in South Asia.5 6

Formative research by CARE’s Tipping Point Initiative (TPI), a programme aimed at addressing the causes of CEFM in Nepal and Bangladesh, has identified CEFM as a symptom of gender inequality rooted in the control of girls’ sexuality and enforced through the exclusion of girls’ voices in marriage processes.7 In Nepal, child marriage...
is concentrated in specific castes—including Dalits, Madhesi, low-caste Hindus and other economically marginalised castes. The social and economic isolation of these groups limits their ability to change practices like child marriage, even when aggregate patterns for all castes suggest change.

Interventions to prevent CEFM

In addition to poor knowledge about causes of CEFM in South Asia, evidence of the impacts of programmes seeking to address CEFM is lacking. In a systematic review of reviews of intervention studies aiming to reduce gender-based violence against adolescent girls, 12 intervention studies cited in high-quality reviews focused on preventing CEFM. Of these, six studies were undertaken in South Asia, and one (quasi-experimental) study was undertaken in Nepal. While a majority of the studies had multiple-component, multilevel intervention designs, almost all of those that involved community-engagement/norms-change activities with individual-level empowerment activities were small and quasi-experimental, limiting the evidence on these types of interventions. An unpublished review of child-marriage intervention studies not included in the above review of reviews revealed similar findings.

CARE’s TPI

CARE’s TPI focuses on addressing the causes of CEFM and on promoting the rights of adolescent girls through community-level programming focusing on a synchronised engagement of different participant groups to challenge social expectations and repressive gender norms and promote girl-centric and girl-led activism. TPI has designed a ‘core’ programme package, the Tipping Point Programme (TPP), which includes components, implemented over 18 months, to enhance adolescent girls’ personal assets/intrinsic agency (including self-efficacy) and instrumental agency (including voice and negotiation skills). TPI also is testing an enhanced programme package, TPP+, also implemented over 18 months, which includes all elements in TPP plus activities to enhance social norms change by engaging community leaders and by facilitating girl-led community activities. Figure 1 summarises the components and participant groups of the CARE ‘core’ and ‘enhanced’ programme models, detailed elsewhere (https://caretippingpoint.org/implementation-manuals/). Programme updates related to COVID-19 are being recorded in the trial registry.

In sum, programmatic gaps to address the causes of CEFM are large, and few, rigorous impact evaluations of social norms and empowerment-based prevention programmes exist. The CARE TPI leveraged its ‘core’ and ‘enhanced’ programme models, a cluster randomised controlled trial (C-RCT), and large sample sizes to assess impacts on CEFM and mediating outcomes aligned with the TPP and TPP+ models.

Figure 1 Care Tipping Point Programme (TPP) and Tipping Point Programme plus intervention packages. ASRHR, Adolescent sexual and reproductive health and rights; G, Girl; BFM, Boy-Father-Mother; SH, Sexual Health; VSLA, Village Savings and Loan Association.

METHODS AND ANALYSIS

Aims and research questions

This study aims to triangulate data from qualitative and quantitative approaches to evaluate the impacts of the TPP and TPP+ models on CEFM by enhancing adolescent girls’ intrinsic, instrumental and collective agency as well as through community social norms diffusion. The following research questions, aligned with the TPI learning components, organise the research. First, to what extent, due to the programme(s), do adolescent girls experience decreases in their (a) risk (qualitative) and (b) hazard (quantitative) of first marriage and first gauna.

Second, to what extent, due to the programme(s), do any observed changes in the above primary outcomes operate through increases in the following secondary outcomes related to girls:

- Personal assets and intrinsic agency, including: (1) critical awareness about gender and rights; (2) aspirations and choices regarding education, freedom of movement and marriage; (3) knowledge about SRHR and (4) attitudes about SRHR.
- Instrumental agency, including (1) communication and negotiation skills; (2) capacity to make decisions, (3) practice of SRHR and (4) leadership competence.
- Collective agency or girl-centred movement building, including: (1) group cohesion, solidarity and mobilisation skills and (2) autonomous engagement with different networks, the community and governmental and non-governmental stakeholders to change social norms and claim their rights.

Finally, to what extent, due to the programme(s), do any observed changes in the above primary outcomes operate through shifts in the following secondary outcomes related to the community:

- Social norms (qualitative and quantitative) with respect to (1) perceptions of what others do in terms of gender, rights (including ASRHR) and CEFM (including dowry); as well as (2) perceptions of what...
others expect them to do in terms of gender, rights (including ASRHR) and CEFM (including dowry).

Setting
The study sites are Kapilvastu and Rupandehi districts, where the Nepali Government has prioritised CARE programming and where no concurrent CARE or other NGO programming related to CEFM have been underway. The districts have the lowest median ages at first marriage in Nepal (See the website: girlsnotbrides.org https://www.girlsnotbrides.org/child-marriage/nepal/). Both districts exhibit low levels of human development. Although life expectancies at birth in Rupandehi are above the national average (68.0 vs 66.6 years), life expectancy in Kapilvastu (61.3 years) falls below the national average. Rupandehi and Kapilvastu lag behind the national averages for per capita income (US$1123 PPP and US$990 vs US$1160 PPP).

Mixed-methods cluster randomised design
This study protocol is adapted from a C-RCT developed by the International Centre for Diarrhoeal Disease Research, Bangladesh (icddr,b), to evaluate TPI in Bangladesh and Nepal. The adapted protocol added sampling procedures and study forms from the Room-to-Read Girls Education Programme (RTR-GEP) impact evaluation,10 modified study forms developed by icddr,b, added the impact-evaluation theory of change (figure 2) and added analysis plans. The adapted study design entails a mixed-methods, three-arm C-RCT. Arm 1 is the control group. Arm 2 is the TPP group. Arm 3 is the TPP+ group. This design allows for measurement of the effects of:
- TPP on primary and secondary outcomes versus the control condition.
- TPP+ on primary and secondary outcomes vs the control condition.
- TPP+ versus TPP, to assess the incremental effects of emphasised social norms change.

The qualitative evaluation involves data collection with adolescent girls, adolescent boys, parents of adolescents and community stakeholders in eight TPP/TPP+clusters through in-depth interviews (IDIs), key informant interviews (KIs) and focus group discussions (FGDs) with participants. IDIs of the participating adolescents will be repeated at endline to examine change over time.

Client and public involvement
Community members from Phase 1 TPI communities were actively involved in designing TPI during the formative research phase.7 This included participating in a learning and design workshop that included discussions on programme content, structure and duration, as well as secondary (or learning) outcomes. In the TPP+ models,
CARE is implementing, adolescents, their parents and community members are engaged in group dialogues, from which ideas for community events are generated. Adolescent girls are identified to serve as activists in their communities. Community members in the study sites also were involved in piloting all study forms. Dissemination of the findings will include events in the study districts and with Nepali officials.

Sample design
Sampling and randomisation of primary sampling units
In Nepal, wards (the lowest governmental administrative unit) were the primary sampling units. Using the size of the resident population from the 2011 Census of Nepal, 27 wards were randomly selected with probability proportionate to size from each of two study districts. Each selected ward within each district was randomly assigned to one of the three study arms (Also allowing for balance in the sample across study districts, for programme implementation purposes.), for 18 wards per study arm (figure 3).

Cluster selection and pretrial household census
Enumerators for the households census mapped selected wards in the field to ensure an accurate household count or to update the count based on information from ward authorities. Enumerators divided large wards (>200 households) into segments of ~200 households and selected one segment randomly. The selection of segments within wards also minimised the extent of physical adjacency of programme and control segments and any spill-over of programme effects into control areas.

In selected clusters (wards or ward segments), enumerators conducted a household census. Census forms were administered to the most knowledgeable woman member, or alternatively to any knowledgeable adult member (online supplemental file 1). Enumerators collected contact information and data on the household’s caste, religion and language(s) spoken. Marital status, age at marriage and years married were recorded for members above age 8 years. These data generated a sampling frame from which eligible participants for groups in clusters randomly assigned to TPP or TPP+ were identified.

Qualitative cluster selection
For the qualitative research, the team chose 8 of the 36 randomly selected clusters assigned to TPP or TPP+ in each district. The Nepal study team selected clusters to ensure diversity with respect to treatment (TPP and TPP+ clusters were equally represented) and with respect to ethnicity, religion and language. Given the lack of ward-level data at the time of cluster selection, CARE Nepal staff shared insights to identify sites that were diverse and thought to have high levels of child marriage. Characteristics of these sites, based on field reports, are shown in table 1.

Eligibility
In treatment and control arms, eligible girls and boys were unmarried, 12–16 years, living in selected clusters, with no plans to migrate in the subsequent 24 months. In clusters randomised to receive TPP or TPP+, participation was voluntary for ethical reasons, and so eligible girls and boys in these clusters also were consenting TPP/TPP+ participants. To measure community norms, eligible adults in treatment and control arms were men and women 25 years or older, living in selected clusters, and with no known plans to migrate in the subsequent 24 months.

For the qualitative research, eligible adolescents were unmarried girls and boys 12–16 years who were living in selected study clusters. Eligible mothers and fathers were male and female parental figures of an adolescent boy or girl who consented to participate in the programme and associated measurement. Eligible community stakeholders were male or female community leaders identified by sponsor organisational contacts in selected clusters.
Survey sample size and power

For the survey, each cluster was intended to have 20 adolescent girls and 20 adolescent boys at endline. Assuming a 50% prevalence of the primary (CEFMs) and secondary (agency) outcomes to allow for maximum sample size, a 5% significance level, 80% power, an intra-cluster correlation of 0.05 (based on prior work in Bangladesh,\textsuperscript{11} and a 15% effect size between study arms, 18 clusters per study arm and 54 clusters total were needed. Assuming a greater than 90% participation rate at baseline and an 85% retention rate at endline for girls and boys, the target baseline enrolment was 23 girls and 23 boys per cluster, or 1242 girls and 1242 boys in total. For the adult community sample, to assess social-norms change, assuming 50% prevalence of CEFM-related norms and a 15% change after the intervention, 5% significance level, 80% power and 5% non-response rate, 90 adult women and 90 adult men 25 years or older were needed per arm, for a total sample size of 540 adult women and men.
Baseline recruitment
From the census data, names and contact information for eligible, randomly selected adolescent girls (N=27) and boys (N=27) in TPP and TPP+ clusters were provided to CARE to form programme groups. Following these randomly ordered lists, CARE Tipping Point project staff recruited and enrolled each eligible adolescent to participate in programming. The local research partner followed up with consenting participants to administer the informed consent and survey until the target number of participants was reached per cluster (23 girls, 23 boys) or all 27 individuals had been approached, whichever came first. Thus, the survey sample represents those who agreed to participate in TPI, and not the population. In control clusters, and for the adult community samples in all clusters, the research team randomly selected eligible participants from cluster-specific lists, invited them to participate in the study, administered the informed consent, and administered the survey to consenting individuals.

For the qualitative research, lists of eligible adolescent girls and boys and their adult care providers were generated from the census data from the eight selected clusters. Lists indicated who had been selected for programme participation to achieve balance in group discussions (see below). Grandparents replaced parents in group discussions when parents had migrated for work.

Achieved samples
Achieved baseline survey sample sizes were 270 adult women, 270 adult men, 1134 unmarried adolescent girls and 1154 unmarried adolescent boys. The total achieved sample sizes across study arms are summarised in figure 4. Sampling for the qualitative research continued until the planned number of IDIs, KIIs and FGDs were completed.

Study forms
Baseline quantitative assessments
The baseline survey in Nepal included 20 modules that measured personal assets/intrinsic agency, instrumental agency, collective agency, social networks and norms, and discrimination and violence as barriers to change (table 2). Nineteen modules were administered to girls; 16 to boys, and 5 to adult community members. Questionnaires were customised for each sample but maintained high comparability (online supplemental file 2).

Endline quantitative assessments
With minimal modifications to ensure comparability, baseline survey forms will be readministered at endline, after the 53-month TPP and TPP+ programme period and 8-month period without programming (to assess sustained impact). To minimise attrition, the endline survey will occur when community residents who migrate for work are expected to return for harvesting. Assessor...
will be blinded to the treatment to which respondents have been exposed. Also, adolescents in programme and control areas will be asked about their exposure to TPP and TPP+ sessions and activities to glean a general understanding of programme fidelity in terms of sessions held as well as potential spill-over of programmatic activities in control clusters.

**Baseline qualitative guides**

Baseline qualitative methods included IDIs with 20 adolescent girls and 10 adolescent boys, 8 KIs with adult community leaders, 16 FGDs with 8 groups of adolescent girls and 8 groups of adolescent boys, and 16 FGDs with parents of adolescents in study communities (Table 3, online supplemental file 3). IDIs with adolescents provided narrative data on attitudes and perceptions towards gender roles, life aspirations, girls’ safety and security, and girls’ mobility. FGDs with adolescents and parents were based on CARE’s social norms analysis plot framework,\(^\text{12}\) highlighting social norms surrounding girls’ mobility, decision making around marriage, and interaction with boys. KIs with teachers and local officials gathered data on collective action, perceived changes in the prevalence of child marriage, social norms and practices around marriage, work and education, and girls’ safety and security.

**Programme monitoring forms**

CARE developed standards for assessing internal implementation fidelity, including for the structure and daily functioning of the programme, to ensure adherence to the intended programme packages and high-quality implementation across programme arms (TPP and TPP+). CARE has assessed implementation fidelity through observations of at least 10% of the sessions every month and feedback from all Tipping Point participants at least every 6 months. These observations and feedback sessions are discussed internally monthly and quarterly to assess the project’s fulfillment of these standards. The research partners also are conducting independent monitoring visits, including observations of the quality of the sessions and skill of the facilitators, adherence to the project’s implementation manuals, and participants’ and communities’ response to TP activities and content. The Emory/IDA team are administering monitoring forms in TPP and TPP+ sites twice, at 6 monthly intervals from baseline during programme implementation (online supplemental file 4). Table 4 summarises the forms, samples and topics covered, including questions to understand the potential impacts of COVID-19 risk mitigation strategies on families in TPP and TPP+ study arms. The 24 session-observations elicit information about the fidelity of facilitators to programme content and delivery. Recommendations provided by the research partners are reviewed in quarterly meetings with field facilitators and higher management of the implementing partners.

**Training**

For the household census, enumerators received a 2-day training (29–30 April 2019) on the study design and aims, use of mobile data collection equipment, and research ethics. The census occurred 1–29 May 2019, visiting 9289 households and documenting 46578 household members.

---

### Table 3 Qualitative methods and samples interviewed, Care Tipping Point Impact Evaluation in Nepal

<table>
<thead>
<tr>
<th>Qualitative field guide</th>
<th>Constructs measured</th>
<th>Samples receiving</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. In-depth interviews</td>
<td>Adolescents: individual perceptions, practices and experiences of gender roles and responsibilities; education/employment aspirations; decision making around marriage; knowledge of ASRHR; girls’ safety/security/mobility</td>
<td>20 10 0 0</td>
</tr>
<tr>
<td>2. Key Informant interviews</td>
<td>Teachers, local government officials: collective action, perceived changes in prevalence of child marriage, social norms and practices around marriage, work and education, girls’ safety/security</td>
<td>0 0 0 8</td>
</tr>
<tr>
<td>3. Focus group discussions</td>
<td>Adolescents: vignettes based on CARE’s SNAP framework, high-lighting social norms about girls’ mobility, marriage decisions</td>
<td>8 8 16 0</td>
</tr>
<tr>
<td>4. Group sizes, range</td>
<td>Parents: social norms diagnosis approach based on CARE’s SNAP framework; girls’ mobility, gender roles/responsibilities, education, interactions with boys</td>
<td>6 6 5–6 –</td>
</tr>
</tbody>
</table>

SNAP: social norms analysis plot.
A 10-day training (30 May 2019–8 June 2019) for baseline included training and practice with study forms, tablet-based data collection and data quality. All data collectors received training on research ethics, including CARE’s reporting protocols surrounding gender-based violence. The training also included field test of the baseline survey forms in one of the two study districts. In addition to the general training, qualitative data collectors received a 2-day training on gender, rights, empowerment, CEFM and qualitative research methods.

**Recruitment and retention**

To recruit and retain participants in the programme and trial, CARE first adapted and tested an intervention package that would engage participants, especially boys and girls. Second, each group agreed on timing and location of meetings with participants to facilitate attendance. Third, a follow-up is triggered for participants who miss two consecutive sessions, for whom Tipping Point facilitators make a home visit to discuss barriers to attendance and to problem-solve re-engagement. Finally, the team plans to implement a two-phase tracing protocol at endline, following guidance on origin-based sampling and phone-tracing techniques from Nepal.13 First, during the endline survey, we will conduct brief, in-person proxy interviews with primary caretakers or household members closest to participants who are lost to in-person follow-up at endline. Second, we will conduct brief, postendline interviews over the phone with those same participants using a shortened questionnaire adapted for phone interviews. The goals of this protocol will be to maximise overall retention of programme participants and to minimise any differential retention across study arms that may arise from differential contact with participants in the programme and control arms.

**Fieldwork**

All data collectors spoke Nepali and the local languages—Awadhi and Bhojpuri—and were gender-matched with participants. Qualitative data collection occurred 7–20 June 2019. Survey data collection occurred 10 June 2019–10 July 2019.

**Data management, access and sharing**

The REDCap mobile app and android tablets permit the secure collection, transfer, and storage of all quantitative study data. Study data are available to trained Emory study staff for cleaning and analysis. Deidentified study data will be stored and shared with collaborating institutions via Emory Box, a password protected, HIPAA-compliant cloud drive. Final study data will be made public by mutual agreement in keeping with CARE and Emory policies regarding data sharing (see ClinicalTrials.gov registry).

**Quantitative data analysis**

**Primary and secondary outcomes**

Table 5 summarises the primary and secondary outcomes and expected impacts of TPP and TPP+ on each outcome. The primary outcome is CEFM during the programme or freeze period, and if so, the age of first marriage.
Secondary outcomes are organised following the theory of change (figure 2). Outcomes are summative scores for measures of personal assets (knowledge about SRH; membership or leadership in groups), intrinsic agency (aspirations about marriage/education/work; self-efficacy; attitudes about gender roles/gender discrimination/menstruation/masculinity), instrumental agency (mobility of girls; communication with parents; leadership competence; participation in financial activities and decision making), collective agency (collective efficacy; collective action), changes in repressive social norms (community social norms), and reductions in violence and discrimination (gender discrimination in the family; public violence/harassment against girls).

Descriptive analyses of pretrial household census data

The Emory team was responsible for cleaning and analysis of pretrial census data. Census data on individual household members were used to estimate the percentages of men and women members, by 5-year age groups, who were married, with and without gauna, before the ages of 15 and 18. Married individual household members reported their ages at marriage and gauna, and these data were used to estimate median ages at marriage and gauna by district, gender and age group (table 6). These census data also will be used after endline to compare, by study arm, the characteristics of sample participants at each wave with those in the census sample who would have been eligible for participation but were not included. Any differences between the census sample and study participants by arm will be reported, and considered in interpretations of analyses that assess TPP and TPP+ impacts.

Scale construction and reliabilities

Individual survey questions (items) for adults, adolescent girls, and adolescent boys will be organised into item sets capturing TPI secondary outcomes. Items will be recoded to be anchored at zero. Missing responses will be coded as missing for univariate analyses of items. Pearson pairwise correlations will ensure that items within sets are mutually correlated and summative scales are reasonable reflections of intended secondary outcomes. An item will be considered for deletion if the magnitude of its pairwise correlation with others in the same item set is close to zero and not significant. Scale reliabilities will be estimated for all outcomes using Cronbach’s alpha for each item set or

<table>
<thead>
<tr>
<th>Table 5</th>
<th>Expected primary and secondary (intermediate) outcomes of the Tipping Point Impact evaluation in Nepal</th>
</tr>
</thead>
<tbody>
<tr>
<td>Primary outcome</td>
<td></td>
</tr>
<tr>
<td>CEFM</td>
<td></td>
</tr>
<tr>
<td>Hazard of marriage from baseline to endline</td>
<td>↓</td>
</tr>
<tr>
<td>Hazard of gauna from baseline to endline</td>
<td>-</td>
</tr>
</tbody>
</table>

| Secondary outcomes | | |
| Personal assets and Intrinsic agency | | |
| Knowledge of sexual and reproductive health and rights (nine items, score 0–27) | - | ↑ | ↑↑ |
| Aspirations about marriage/education/work (frequencies) | - | ↑ | ↑↑ |
| Self-efficacy (11 items, score 0–22; 13 items for girls, score 0–28) | - | ↑ | ↑↑ |
| Attitudes about gender/gender discrimination/menstruation/masculinity (nine items, score 0–27) | - | ↑ | ↑↑ |

| Instrumental agency | | |
| Mobility of girls (nine items, score 0–18) | - | ↑ | ↑↑ |
| Communication with parents (six items, score 0–18) | - | ↑ | ↑↑ |
| Leadership competence (nine items, score 0–27) | - | ↑ | ↑↑ |
| Participation in financial activities (frequencies) | - | ↑ | ↑↑ |
| Participation in decision-making (frequencies) | - | ↑ | ↑↑ |

| Girl-centred movement building and collective agency | | |
| Collective efficacy (nine items, score 0–15) | - | ↑ | ↑↑ |
| Collective action | - | ↑ | ↑↑ |

| Changes in repressive social norms | | |
| Community social norms (11 items for norms about women and girls, score 0–33; five items for norms about boys, score 0–15) | - | ↑ | ↑↑ |

| Reductions in violence and harassment as barriers to change | | |
| Gender discrimination in the family (five items, score 0–15) | - | ↑ | ↑↑ |
| Public violence/harassment against girls | - | ↑ | ↑↑ |

Control communities are expected to show little change over the study period due to the concentration of child marriage in these conservative communities and the period of observation. The symbol ↑ refers to an increase, and the symbol ↑↑ refers to a more substantial increase. The symbol ↓ refers to a decrease, and the symbol ↓↓ refers to a more substantial decrease.

CEFM, child, early and forced marriage; TPP, Tipping Point Programme.
## Table 6  
Comparison of marital status, age and marriage and gauna, and percentages of household members married by ages 15 and 18, by study arm, pretrial census data for 54 Clusters, Kapilvastu and Rupandehi districts, Nepal

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Women, N</td>
<td>1439</td>
<td>1226</td>
<td>1109</td>
<td>910</td>
<td>769</td>
<td>745</td>
<td>640</td>
<td>508</td>
<td>3572</td>
<td>1586</td>
</tr>
<tr>
<td>Married by age 15, %</td>
<td>–</td>
<td>5.6</td>
<td>16.6</td>
<td>28.2</td>
<td>39.9</td>
<td>47.3</td>
<td>51.9</td>
<td>56.3</td>
<td>43.0</td>
<td>63.1</td>
</tr>
<tr>
<td>Married by age 18, %</td>
<td>–</td>
<td>–</td>
<td>40.9</td>
<td>51.7</td>
<td>68.9</td>
<td>76.2</td>
<td>76.6</td>
<td>79.1</td>
<td>68.9</td>
<td>81.1</td>
</tr>
<tr>
<td>Median age at first marriage of those married</td>
<td>13.0</td>
<td>16.0</td>
<td>17.1</td>
<td>17.1</td>
<td>16.2</td>
<td>15.2</td>
<td>15.0</td>
<td>14.8</td>
<td>15.8</td>
<td>13.6</td>
</tr>
<tr>
<td>Median age at gauna of those married</td>
<td>–</td>
<td>16.7</td>
<td>18.2</td>
<td>18.4</td>
<td>17.7</td>
<td>17.3</td>
<td>17.0</td>
<td>16.8</td>
<td>17.5</td>
<td>15.7</td>
</tr>
<tr>
<td>Men, N</td>
<td>1593</td>
<td>1245</td>
<td>934</td>
<td>770</td>
<td>599</td>
<td>654</td>
<td>535</td>
<td>521</td>
<td>3079</td>
<td>1695</td>
</tr>
<tr>
<td>Married by age 15, %</td>
<td>–</td>
<td>1.1</td>
<td>6.6</td>
<td>13.4</td>
<td>19.9</td>
<td>29.1</td>
<td>33.5</td>
<td>37.6</td>
<td>25.6</td>
<td>45.4</td>
</tr>
<tr>
<td>Married by age 18, %</td>
<td>–</td>
<td>–</td>
<td>18.0</td>
<td>27.4</td>
<td>40.4</td>
<td>50.3</td>
<td>53.1</td>
<td>54.5</td>
<td>43.9</td>
<td>60.4</td>
</tr>
<tr>
<td>Median age at first marriage of those married</td>
<td>16.6</td>
<td>18.4</td>
<td>19.3</td>
<td>19.7</td>
<td>18.3</td>
<td>17.7</td>
<td>17.7</td>
<td>18.6</td>
<td>16.9</td>
<td>16.9</td>
</tr>
<tr>
<td>Median age at gauna of those married</td>
<td>–</td>
<td>17.3</td>
<td>19.3</td>
<td>20.6</td>
<td>21.2</td>
<td>20.2</td>
<td>20.0</td>
<td>20.3</td>
<td>19.4</td>
<td>19.5</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Women, N</td>
<td>1587</td>
<td>1297</td>
<td>1350</td>
<td>970</td>
<td>751</td>
<td>757</td>
<td>525</td>
<td>458</td>
<td>3461</td>
<td>1489</td>
</tr>
<tr>
<td>Married by age 15, %</td>
<td>–</td>
<td>7.6</td>
<td>25.4</td>
<td>45.7</td>
<td>59.9</td>
<td>62.4</td>
<td>68.8</td>
<td>67.0</td>
<td>58.7</td>
<td>72.9</td>
</tr>
<tr>
<td>Married by age 18, %</td>
<td>–</td>
<td>–</td>
<td>52.2</td>
<td>68.4</td>
<td>79.8</td>
<td>85.9</td>
<td>86.3</td>
<td>85.6</td>
<td>79.7</td>
<td>85.0</td>
</tr>
<tr>
<td>Median age at first marriage of those married</td>
<td>15.8</td>
<td>15.8</td>
<td>16.3</td>
<td>15.6</td>
<td>14.1</td>
<td>13.9</td>
<td>13.0</td>
<td>13.3</td>
<td>14.2</td>
<td>12.3</td>
</tr>
<tr>
<td>Median age at gauna of those married</td>
<td>–</td>
<td>16.7</td>
<td>17.9</td>
<td>18.0</td>
<td>17.0</td>
<td>16.9</td>
<td>16.4</td>
<td>16.8</td>
<td>17.1</td>
<td>16.4</td>
</tr>
<tr>
<td>Men, N</td>
<td>1824</td>
<td>1324</td>
<td>1082</td>
<td>842</td>
<td>668</td>
<td>727</td>
<td>510</td>
<td>476</td>
<td>3223</td>
<td>1698</td>
</tr>
<tr>
<td>Married by age 15, %</td>
<td>–</td>
<td>1.3</td>
<td>7.8</td>
<td>19.5</td>
<td>32.0</td>
<td>39.2</td>
<td>42.4</td>
<td>49.2</td>
<td>34.5</td>
<td>54.8</td>
</tr>
<tr>
<td>Married by age 18, %</td>
<td>–</td>
<td>–</td>
<td>20.6</td>
<td>79.8</td>
<td>54.5</td>
<td>62.0</td>
<td>61.6</td>
<td>65.1</td>
<td>54.2</td>
<td>69.3</td>
</tr>
<tr>
<td>Median age at first marriage of those married</td>
<td>8.0</td>
<td>16.9</td>
<td>18.1</td>
<td>18.5</td>
<td>17.3</td>
<td>16.7</td>
<td>16.0</td>
<td>17.0</td>
<td>15.2</td>
<td>15.2</td>
</tr>
<tr>
<td>Median age at gauna of those married</td>
<td>–</td>
<td>13.0</td>
<td>17.3</td>
<td>19.4</td>
<td>20.6</td>
<td>19.7</td>
<td>19.1</td>
<td>19.6</td>
<td>19.8</td>
<td>19.2</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Women, N</td>
<td>1419</td>
<td>1068</td>
<td>996</td>
<td>769</td>
<td>674</td>
<td>661</td>
<td>531</td>
<td>434</td>
<td>3069</td>
<td>1346</td>
</tr>
<tr>
<td>Married by age 15, %</td>
<td>–</td>
<td>5.8</td>
<td>18.9</td>
<td>32.0</td>
<td>40.5</td>
<td>52.8</td>
<td>48.8</td>
<td>55.1</td>
<td>44.5</td>
<td>61.6</td>
</tr>
<tr>
<td>Married by age 18, %</td>
<td>–</td>
<td>–</td>
<td>42.6</td>
<td>56.7</td>
<td>67.2</td>
<td>76.4</td>
<td>73.3</td>
<td>76.0</td>
<td>68.9</td>
<td>78.5</td>
</tr>
<tr>
<td>Median age at first marriage of those married</td>
<td>15.7</td>
<td>17.0</td>
<td>16.4</td>
<td>15.8</td>
<td>15.0</td>
<td>15.1</td>
<td>14.7</td>
<td>15.5</td>
<td>13.5</td>
<td>13.5</td>
</tr>
<tr>
<td>Median age at gauna of those married</td>
<td>16.5</td>
<td>18.2</td>
<td>17.8</td>
<td>17.5</td>
<td>17.3</td>
<td>17.2</td>
<td>17.2</td>
<td>17.4</td>
<td>15.9</td>
<td>15.9</td>
</tr>
<tr>
<td>Men, N</td>
<td>1437</td>
<td>1082</td>
<td>833</td>
<td>647</td>
<td>537</td>
<td>588</td>
<td>492</td>
<td>464</td>
<td>2728</td>
<td>1527</td>
</tr>
<tr>
<td>Married by age 15, %</td>
<td>–</td>
<td>0.9</td>
<td>6.7</td>
<td>15.3</td>
<td>20.7</td>
<td>28.7</td>
<td>29.1</td>
<td>34.1</td>
<td>24.9</td>
<td>42.1</td>
</tr>
<tr>
<td>Married by age 18, %</td>
<td>–</td>
<td>–</td>
<td>19.3</td>
<td>33.5</td>
<td>41.2</td>
<td>52.9</td>
<td>50.2</td>
<td>55.4</td>
<td>45.9</td>
<td>58.9</td>
</tr>
<tr>
<td>Median age at first marriage of those married</td>
<td>15.6</td>
<td>18.2</td>
<td>19.0</td>
<td>18.6</td>
<td>18.0</td>
<td>17.9</td>
<td>17.3</td>
<td>18.2</td>
<td>16.6</td>
<td>16.6</td>
</tr>
<tr>
<td>Median age at gauna of those married</td>
<td>–</td>
<td>16.8</td>
<td>19.3</td>
<td>20.3</td>
<td>20.2</td>
<td>20.0</td>
<td>20.1</td>
<td>19.7</td>
<td>20.1</td>
<td>19.3</td>
</tr>
</tbody>
</table>

TPP, Tipping Point Programme.
about potential non-arms

Addressing observed and unobserved imbalances across study group, in a full mediation model, following figure 
assess the impacts of TPP and TPP+, relative to the control 
are observed, we will use propensity score methods18–20 
observed characteristics across study arms. If imbalances 
same study arm who would have been eligible and report 
will compare characteristics of TPP/TPP+ participants 
study arms that may be introduced by unintended failure 
We will include covariates in a probit model to predict 
voluntariness of participation for ethical reasons.18 
preprogramme differences in participation resulting from 
bias in the estimated treatment effect due to observed 
given a set of covariates. This method aims to reduce overt 
covariates. For secondary outcomes, we will estimate (non- 
for example, district strata, observed variables that are 
unbalanced across study arms at baseline, and potential 
question whether scales for the secondary 
outcomes are unidimensional or multidimensional.

Baseline descriptive analysis of secondary outcomes and 
covariates
Mean scores, and score tertiles for secondary outcomes 
will be reported. Demographic characteristics and 
scales for secondary outcomes will be compared across 
study arms (control, TPP, TPP+) to assess within-sample 
balance across study arms. Group differences will be 
assessed using χ² tests of independence for categorical 
variables and t-tests for continuous (or quasi-continuous) 
outcomes, accounting for the cluster-sample design. All 
descriptive analyses will be completed in Stata V.16.

Time-to-event and intention-to-treat impact assessment
For the CEFM outcomes, given that the data are right 
censored, we will estimate marginal Cox proportional 
ods models14 15 to estimate the impact of the treatment 
groups on the hazard of marriage or gauna in months 
from baseline, relative to the control group, adjusted for 
covariates. For secondary outcomes, we will estimate (non- 
parametrically) the average treatment effect (ATE) by 
computing the difference between their means for inter-
vention participants and non-participants. We will use the 
difference-in-difference (DiD) regression approach with 
cluster-robust variance estimators.16 We will perform a 
robustness check of the DiD results using the predicted 
hazard of marriage or gauna for each secondary outcome as 
sampling weights.17 The DiD model may be extended to 
assess the impacts of TPP and TPP+, relative to the control 
group, in a full mediation model, following figure 2.

Addressing observed and unobserved imbalances across study 
arms
In the above impact assessments, we address questions 
about potential non-comparability of the samples across 
study arms that may be introduced by unintended failure 
of the randomisation process and/or the voluntary 
participation of adolescents in TPP and TPP+. First, we 
will compare characteristics of TPP/TPP+ participants 
at each wave with those from the census sample in the 
same study arm who would have been eligible and report 
any differences. Second, we assess ‘balance’ on a range of 
observed characteristics across study arms. If imbalances 
are observed, we will use propensity score methods18–20 
that team members have used previously21 to generate 
conditional probabilities of enrolment in TPP and TPP+, 
given a set of covariates. This method aims to reduce overt 
biases in the estimated treatment effect due to observed 
proprogramme differences in participation resulting from 
the voluntariness of participation for ethical reasons.18 
We will include covariates in a probit model to predict 
the likelihood of programme participation, including,
household socioeconomic conditions, employment status of household head, household size, etc).

Sensitivity analysis on the treatment effect in the presence of non-compliance and contamination

We will assess the robustness of the inferences regarding the treatment effects using an instrumental variable approach with the randomisation as the instrumental variable.25

Qualitative data analysis

Baseline qualitative data analysis

Emory’s team revised a qualitative codebook developed by icddr,b for Bangladesh26 to streamline existing codes and add new codes on emergent themes from the Nepal data (online supplemental file 5). Several rounds of codebook edits were made, first through careful reading of 10 transcripts and discussions with CARE and icddr,b; second, after two rounds of intercoder reliability testing (evaluated using Cohen’s kappa27 among three Emory team members using seven transcripts, and finally, after coding 20 transcripts across two sites. Team debriefs were used to resolve discrepancies and make minor edits to codes and definitions, after which the same three-member team divided the remaining 50 transcripts to code individually. All coding, cross-classification and inter-coder reliability testing was performed in MAXQDA V.18 (Berlin, Germany) by this team.

A narrative analysis of IDIs with adolescents began with memos of each transcript summarising themes of ASRH, aspirations, marriage, mobility, safety and security. Descriptive analysis of social norms data across all 70 transcripts was performed by crossing each major theme with relevant norms codes (normative expectations, empirical expectations, sanctions/sensitivity to sanctions, exceptions). Thick descriptions were generated for five norms CARE identified through formative research.26 Documentation of analytical process through memos and triangulation across samples and data collection methods permitted validation at each analytic phase.

Longitudinal qualitative data analysis

We will repeat the above process with endline data and conduct structured comparisons between baseline and endline data within themes as well as by gender, site, caste/community and any other emergent variable found to be relevant to the change process. Finally, we will generate thick descriptions for all codes representing the theory of change and apply causal chain analysis28 to interrogate the components of the theory of change.

DISCUSSION

Limitations and strengths of the TPI impact evaluation design in Nepal

► Despite randomising wards to study arms, samples may differ on selected characteristics (school type, caste, religion, girls’ literacy), which we will control for in analyses.

► Field staff reported challenges collecting accurate data on age to determine eligibility, which the team addressed with repeated verification. Data on age were first gathered during the household enumeration exercise by IDA-trained staff and data collectors. This information was then verified, and inconsistencies resolved, while CARE and implementing partner staff were forming groups of eligible programme participants to ensure criteria for intervention inclusion were met.

► As with age, accurate estimates of age at marriage required repeat verification and triangulation during data cleaning and analysis. Triangulation of age and age of marriage during data cleaning and analysis was achieved by including questions in the enumeration form (current age, age at marriage, years married) that allowed for consistency checks in age reporting and by assessing concordance in years married with husband–wife dyads.

► The Nepal TPI study team faced challenges recruiting participants for boys’ and girls’ groups in the intervention study arms, so the team sought informed consent from and surveyed only those who agreed to participate in TPI/TPP+. This decision may have resulted in potentially non-representative samples of boys’ and girls’ in intervention arms, differences in observed and unobserved characteristics of adolescents in treatment and control clusters, the inability to construct sampling weights, and the need to generalise findings to the study sample. Despite this caveat, following a random probability sample of adolescents in the control clusters to understand their trajectories in primary and secondary outcomes in the absence of TPP/TPP+ is important. Also, we have proposed robust methods to diagnose differences between programme participants and otherwise eligible household census members who did not participate. We also have proposed methods to address systematic and meaningful imbalances in observed characteristics across treatment and control arms and to test for the influence of unobserved characteristics (endogeneity) in ATEs. We will interpret these findings judiciously in light of the full set of findings.

► Differential attrition across study arms may arise at endline from the intensive follow-up of TPP and TPP+ participants during the 18 months of programming after baseline and the 30 months of non-contact with participants in the control arm after baseline. After endline, when the extent of attrition is known for all samples in all study arms, we will implement a targeted tracing protocol to mitigate differences in attrition. Despite these limitations, the TPI Impact Evaluation in Nepal has notable strengths.

► The mixed-methods design allows for triangulation of qualitative and quantitative findings.
The three-arm C-RCT design permits assessment of the incremental impacts of the emphasised social-norm change component.

Robust field strategies and analytical strategies are proposed to mitigate identified risks and caveats in the study design.

Implications of the study for programme implementation and scale-up

The CARE TPI is a novel, integrated, social norms and girl-led movement-building initiative that is designed to address the causes of CEFM through dialogue with girls, boys and community members. The initiative is being undertaken in communities of Nepal with high levels of child marriage despite moderate standards of living. The impact evaluation integrates rigorous qualitative and quantitative methods in a cluster randomised controlled design, allowing for inferences to be made within a sample that is intended to be broadly similar to the population of the two study districts, suggesting that the trial may offer insights to the population in these areas. If found to be effective in the study area, the TPI will have enormous potential for national scale-up as well as adaptation to other settings where CEFM remains prevalent.

Ethics and dissemination

CARE Nepal has ongoing approval from the Social Welfare Council of the Nepali government to operate in the study districts. The Emory University Institutional Review Board (IRB00109419) and Nepal Health Research Council (161 2019) approved the study. The team is following UNICEF guidelines for ethical research involving children, which corroborate standard ethical guidance, and ensure a child-centred focus. CARE Guidelines for Interviewing Children and gender-based violence communications guidelines also are being followed. Written consent forms for adults, as well as parental consent and child assent forms for children (online supplemental file 6), were developed following those from RTR-GEP impact evaluation in Nepal. The data from this study are available from the corresponding author on reasonable request. All other research materials are included in this published article and its online supplemental information files.

Author affiliations

1 Global Health & Sociology, Emory University School of Public Health, Atlanta, Georgia, USA
2 Global Health, Emory University School of Public Health, Atlanta, Georgia, USA
3 Psychology, Emory University, Atlanta, Georgia, USA
4 Gender Justice Team, CARE USA, Atlanta, Georgia, USA
5 Interdisciplinary Analysts (IDA), Kathmandu, Nepal
6 International Centre for Diarrhoeal Disease Research Bangladesh, Dhaka, Bangladesh
7 Health Equity and Rights Team, CARE USA, Atlanta, Georgia, USA

Twitter Cari Jo Clark @cari_jo_clark

Collaborators

Tipping Point Team, CARE: Shikha Sunuwar, Rajan Subedi, Dipendra Raj Sharma, Tirzah Brown; Interdisciplinary Analysts (IDA): Mr. Hiranya Baral, Ms. Sveetcha Ghimire, Mr. Digvijay Mishra, Mr. Sandeep Thapa, Mr. Santosh Kumar

Karki, Ms. Akriti Rana, Mr. Pankaj Pokhrel, Ms. Prakriti Adhikary, Mr. Tikaram Basnet, Ms. Nishu Aryal, Mr. Ram Ishwor Yadav, Ms. Mamta Hamal, Ms. Sarala Regmi, Mr. Anil Chaudhary, Ms. Dilmya Dhakal, Mr. Gajendra Prasad Sah, Ms. Barsha Glen, and Mr. Nischal Rai Dawadi.

Contributors

KMY: conceptualisation, data curation, funding acquisition, investigation, methodology, project administration, resources, software, supervision, visualisation, writing-original draft, writing-review and editing. IB: data curation, project administration, writing-original draft, writing-review and editing. YFC: methodology, writing-original draft, writing-review and editing. CC: conceptualisation, data curation, funding acquisition, investigation, methodology, supervision, writing-original draft, writing-review and editing. SK: conceptualisation, funding acquisition, project administration, writing-review and editing. ZK: formal analysis, investigation, writing-original draft, writing-review and editing. AL: supervision, writing-original draft, writing-review and editing. MAM: conceptualisation, methodology, writing-review and editing. RTN: conceptualisation, methodology, writing-review and editing. KP: conceptualisation, methodology, writing-review and editing. AT: conceptualisation, methodology, writing-review and editing. SS: methodology, validation, investigation, resources, data curation, writing-review and editing, supervision. SD: methodology, validation, investigation, resources, data curation, project administration. AS: conceptualisation, funding acquisition, project administration, supervision, writing-review and editing. All authors have read and approved the manuscript.

Funding

The Nepal Tipping Point Trial is supported through a research grant from the Kendeda Fund to CARE USA and subawarded to Emory University.

Disclaimer

The funder played no role in the design of the study; nor in the collection, analysis, or interpretation of the data; nor in the writing of the manuscript.

Competing interests

None declared.

Patient consent for publication

Not required.

Provenance and peer review

Not commissioned; externally peer reviewed.

Supplemental material

This content has been supplied by the author(s). It has not been vetted by BMJ Publishing Group Limited (BMJ) and may not have been peer-reviewed. Any opinions or recommendations discussed are solely those of the author(s) and are not endorsed by BMJ. BMJ disclaims all liability and responsibility arising from any reliance placed on the content. Where the content includes any translated material, BMJ does not warrant the accuracy and reliability of the translations (including but not limited to local regulations, clinical guidelines, terminology, drug names and drug dosages), and is not responsible for any error and/or omissions arising from translation and adaptation or otherwise.

Open access

This is an open access article distributed in accordance with the Creative Commons Attribution Non Commercial (CC BY-NC 4.0) license, which permits others to distribute, remix, adapt, build upon this work non-commercially, and license their derivative works on different terms, provided the original work is properly cited, appropriate credit is given, any changes made indicated, and the use is non-commercial. See: http://creativecommons.org/licenses/by-nc/4.0/.

ORCID iD

Kathryn M Yount http://orcid.org/0000-0003-1917-1574

REFERENCES


11 Parvin K, Nunna T, Mamun M. *Tipping point project: report of the baseline study findings from Bangladesh*. kdcdr, b and CARE USA, 2020.


26 CARE USA. *The cultural context of Child Marriage in Nepal and Bangladesh- Findings from CARE’s Tipping Point Project community Participatory Analysis: Cooperative for Assistance and Relief Everywhere, Inc*, 2016.


33 Bergenfeldt I, Clark CJ, Kalra S. *Tipping point program impact evaluation: baseline study findings in Nepal*. Atlanta, GA: Care USA and Emory University, 2019.