Community-Based Cluster Randomized Controlled Trial: Empowering Households to Identify and Provide Appropriate Care for Low-Birthweight Newborns in Nepal

Stephen Hodgins (shodgins@ualberta.ca)
University of Alberta  https://orcid.org/0000-0001-8365-3311

Binamra Rajbhandari
International Rescue Committee

Deepak Joshi
Save the Children

Bharat Ban
consultant

Subarna Khatry
consultant

Luke Mullany
Johns Hopkins University

Research article

Keywords: low birth-weight, thermal care, care-seeking, anthropometry, home childbirth, behavior change, sensitivity and specificity, low- and middle-income countries

Posted Date: March 5th, 2020

DOI: https://doi.org/10.21203/rs.3.rs-15575/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License. Read Full License

Version of Record: A version of this preprint was published on August 24th, 2020. See the published version at https://doi.org/10.1186/s12889-020-09317-w.
Abstract

Background: Most newborn deaths occur among those of low birthweight (LBWt), due to prematurity &/or impaired fetal growth. Simple practices can mitigate this risk. In low-income country settings where many births occur at home, strategies are needed that empower mothers to determine if their babies are higher risk and take protective measures. Earlier studies suggest that foot-length may be a good proxy for birthweight. An earlier Nepal study found a 6.9cm cut-off performed relatively well, differentiating normal from low birthweight.

Methods: Community-based, cluster-randomized controlled trial. Objective: to determine whether family-administered screening, with targeted messages improves care practices known to mitigate LBWt risks. Participants: pregnant women participating in a parent trial in rural Nepal. Women were given a 6.9cm card to assess whether the baby’s foot is small; if so, to call number on the card for advice. Follow-up visits were made over 2 weeks following the birth, assessing for: reported skin-to-skin thermal care, and care-seeking outside the home; restricting to low birthweight (using 2 cutoffs: 2,500g and 2,000g).

Randomization: 17 clusters intervention, 17 control. The study also documented steps along the presumed causal chain from intervention through behavioral impact.

Results: 2,022 into intervention, 2,432 into control. Intervention arm: 519 with birthweight <2,500g (vs. 663 among controls), of which 503 were available for analysis (vs. 649 among controls). No significant difference on care-seeking; for those <2,500g RR 1.13 (95%CI: 0.97-1.131). More of those in the intervention arm reported skin-to-skin thermal care than among controls; for those <2,500g RR 2.50 (95%CI: 2.01-3.1). Process measures suggest this apparent effect cannot be attributed to the intervention; the card performed poorly as a proxy for LBWt, misclassifying 84.5% of those <2,000 as normal.

Conclusions: Although the trial found an apparent effect on one key behavioral outcome, this cannot be attributed to the intervention; most likely it was a result of pure chance. Other approaches are needed for identifying at-risk babies in such settings, and targeting them for appropriate care messaging.

Trial registration: clinicaltrials.gov identifier: NCT02802332, registered 16 June 2016, https://clinicaltrials.gov/ct2/show/NCT02802332

Background

Globally, it is estimated that approximately 15% of newborns weigh < 2,500 g at birth; rates are the highest in South Asia, where low-birthweight newborns make up over a quarter of births [1]. In Nepal—the setting for the study reported in this paper—32% of newborns were found to weigh < 2,500 g, in a rural terai (plains) population [2]. Most newborn deaths are among such babies of low birth-weight, either those growth-restricted in utero or born preterm; globally, although “low-birthweight babies constitute only about 14% of children born, they account for 60–80% of neonatal deaths” [3].
In principle, many deaths in small newborns could be prevented with good attention to thermal care [4] and breastfeeding [5] (early initiation, exclusive, at adequate frequency), and with prompt medical attention in case of complications. For institutional births, while mother and newborn are still in hospital there is an opportunity to identify such higher-risk newborns, provide them with any needed special care, and make suitable arrangements for follow-up after discharge (though frequently in low-resource settings these newborns do not receive such care). For home births, however, babies at high risk due to low birthweight are often not recognized as such and, therefore, may not receive needed care. In a Nepal setting, it has been found that mother’s judgement of relative size of their newborns is generally not reliable [6]. From the most recent Nepal Demographic and Health Survey [7], 57% of births were in health facilities, mainly government hospitals. However, there were marked urban-rural differences (69% and 44%, respectively) and disparities by wealth quintile, with only a third of those in the bottom quintile giving birth in a health facility.

Nepal benefits from a comparatively robust peripheral-level primary healthcare system, consisting of health posts, each serving populations of 5–10,000, which are staffed by 3 or more fulltime, paid health auxiliaries, who are supported by a network of 9 or more Female Community Health Volunteers (FCHV), most of whom are actively involved in program work—including maternal-child activities—notably advice and support to pregnant women. In a recent, nationally-representative survey of women who had given birth over the preceding year [8], 55% of such women reported having received such support during their last pregnancy. Through the period 2009–2014, there had been program efforts in Nepal, encouraging FCHVs to make early postnatal home visits (part of the intention being to identify and support higher-risk newborns), however—despite promising evidence during piloting—when implementing at scale, coverage was very low [9].

Based on the principle of the household production of health [10], we hypothesized that there may be ways to empower mothers and other household care providers— particularly in instances where births happen at home—to determine, themselves, if their babies are particularly small and, if so, to take appropriate actions to reduce risk. Earlier work in Nepal [2] documented a reasonably good correlation between birthweight and several other anthropometric measures, notably foot-length. Specifically, with a 6.9 cm cut-off, this offered 87.5% sensitivity for a weight threshold of 2,000 g, and specificity of 94.2% for a weight threshold of 2,500 g. In other words, using 6.9 cm as a cut-off, only 12.5% of those weighing < 2,000 g would be falsely classified as normal and, of those > 2,500 g, only 5.8% would be classified as small, potentially needing special care. A dedicated measuring box was used for that study (see Fig. 1).

We speculated that if pregnant women were given a card measuring 6.9 cm, with instructions on its use, they would be able to determine with reasonable accuracy whether their baby was at particular risk due to low birthweight. For such cases, we could target appropriate care messages.

Similar ideas have occurred to other investigators [11, 12], and a variety of strategies have been tested, most often involving issuing a ruler or other measuring device or materials to community health workers, who were then expected to make early postnatal home visits to identify small babies. Although the
Ministry of Health and partners in Nepal have attempted a similar strategy (with the community health volunteers using Salter scales to identify small babies), as mentioned above this hadn't worked well (not only was coverage low but even when first piloted, many FCHVs had difficulty accurately weighing newborns). We were—instead—looking for an approach that would empower family members, themselves, to determine if their babies were particularly small.

Experiences elsewhere are of some relevance. Marchant and colleagues [11] used such an approach in a pilot in southern Tanzania. Community health workers (CHW) were given a laminated counseling card with an integrated measurement area on the bottom right corner of the card—as below (Fig. 2)—to classify newborns by foot-length as: very short (< 7.0 cm), short (7.0-7.9 cm), or not short (≥ 8.0 cm); and the CHWs were also instructed to measure and record foot-length, using a transparent plastic ruler.

In a Nepal-based study in a tertiary-level hospital [13], study staff did measures of newborn foot-length using 3 different methods: 1) a transparent plastic ruler placed against the sole of the foot; 2) foot placed vertically onto a hard wooden surface and then measured using a tape; and 3) foot placed on white paper on a firm surface, tracing the outline with a pencil—with length measured from heel to tip of the big toe in each of the 3 methods. There was some variation in results, with method 2 yielding measures on average 0.2cm longer than for method 3, and measures from method 1 intermediate between the other 2.

Measures using method 1 performed best with regard to receiver-operator curve measures for birthweight; for a birthweight cutoff of 2,000g, a 7.5cm threshold performed best, with sensitivity of 83% and specificity of 85%.

In a more recent study [12], conducted in a tertiary-level hospital in India, the investigators—having concerns about the accuracy of instruments like plastic rulers—instead used a specially-made caliper (see Figure 3). The intention was that such a measuring device and procedure could be used by community-based frontline health workers. Similar to the Nepal study cited above [2], Pratinidhi et al. [12] found that a cut-off of 6.8cm performed best to differentiate by birthweight corresponding to survival risk. With this cut-off, 4% of those weighing <2,000g (n=23) were misclassified as normal weight and only 3.6% of those with birthweight ≥2,500g (n=83) were misclassified as small.

The objective of our study was to determine if an intervention entailing provision to women late in pregnancy of a simple tool allowing them to classify their newborns as small increases uptake of key care practices (mediated through being exposed to recorded messages, heard on calling a toll-free number), notably: skin-to-skin thermal care and seeking treatment for possible danger signs. Implementation of the study as a cluster trial replicates conditions that would be present if this were delivered as a community-based program, in which one might expect amplification of effect through women sharing this information with friends and neighbors.

**Methods**
The study was piggy-backed on a parent cluster-randomized controlled trial implemented in Sarlahi District (in the plains area of Nepal, close to the Indian border), investigating the efficacy of sunflower-seed oil massage on newborn survival [14]. Thirty-four participating government administrative areas (Village Development Committees (VDCs)) were randomized, with 17 allocated to receive the foot-length card/ phone-message intervention and 17 to control. Variables on which the randomization was restricted included:

- Size of birth cohort (i.e. number of live births),
- Low birth weight (i.e. proportion of babies weighing < 2500 grams),
- Birth location (proportion born in facility),
- Skin-to-skin care (proportion of mothers reporting such care in the first week of life), and
- Care seeking (proportion of mothers reporting seeking care for infants in the first week of life).

Tolerance for determining “balance” of the five variables was set at 10%, i.e. a randomization was considered a candidate sequence if the ratio, r, of the variable of interest—when comparing the intervention (foot length card) to no intervention (absence of foot length card)—met the following criteria: 1/1.1 < r < 1.1. Among one million randomization sequences generated, 45,427 sequences met the above criteria, from which a single sequence was randomly selected.

All women in late pregnancy, aged 15 years and older, living in the 34 participating VDCs, and enrolled in the parent trial were eligible to participate in this study. The design of the parent trial and the infrastructure in place for implementing it, meant that field staff already had contact with women in late pregnancy, which enabled identification of eligible participants for this study.

In the intervention arm, during scheduled late-pregnancy visits, study field staff gave the women the cards (see Figure 4) with basic verbal instructions in Maithili or Nepali on their use, depending on their mother language (and the same information in written form). Mothers were told to lay their baby on a firm surface and place the card against the bottom of the baby’s foot to determine whether or not the foot was longer than the card and, if not, to call the toll-free number for helpful care instructions. On calling one of those numbers, the mothers or family members were able to listen to a 2-minute recorded message, in Maithili or Nepali depending on which card they received (see full text in English in the annex).

The recording was intended to conveyed 3 cognitive and 2 affective messages, notably, cognitive: 1) avoiding allowing the baby to get chilled (best to maintain the baby skin-to-skin against the mother’s chest); 2) on feeding - ensuring adequate frequency (at least every 2.5 hours around the clock) and seeking assistance if the baby had difficulty feeding; and 3) recognizing potential danger signs and promptly seeking medical care; and affective: 1) urgency/ importance, and 2) self-efficacy (you can do this, and it will protect your baby).

In both intervention and control clusters, study field-staff made repeated follow-up home visits: within 24 hours of birth (for health facility births, the first postnatal home visit could be on day 3 or later) and then
on days 3, 7, 10, and 14. On each occasion, questions were asked about our principal endpoints (skin-to-skin care and treatment seeking).

In addition, during the first visit after birth (in most instances, within 24 hours of birth), in intervention clusters field research staff inquired concerning card use. Among those reporting having used the card, field staff inquired concerning how the baby was classified (small/ not small), and whether or not caregivers called the toll-free number, and (for those reporting having made the call) they inquired concerning recall of the messages.

Study staff also used the card themselves, and classified the baby’s as having normal or short foot-length. In both intervention and control clusters, using instruments from the parent trial, mothers were queried about care-seeking behavioral outcomes (i.e. any skin-to-skin kangaroo mother care, and any care-seeking from qualified health workers). From data gathered up to 14 days of life, it was determined whether or not either of the target care behaviors were reported.

For sample-size determination, based on earlier data collected in this setting over the previous 4 years, we assumed that 30% of newborns would have birthweight <2500g, and that 30% of those in the non-intervention arm would report seeking care from a health worker and 25%, any skin-to-skin care. Using these inputs, to detect an absolute difference between intervention and control condition of 10 percentage points, with study power of 80%, a minimum of 1,180 newborns would be required in each study arm. To ensure adequate power for our main endpoints and to enable us to investigate process steps, we aimed to enroll at least 2,000 newborns per study arm.

Baseline characteristics of the intervention and control groups were compared to examine the effectiveness of the randomization process and identify any potential confounding factors.

Analysis of study results consisted of two main components:

- Assessing the impact of a communications intervention on two care behaviors with a potential to reduce risk for low birthweight newborns, and
- Process documentation along the hypothesized causal pathway, expected to produce the intervention effect: received card → measured baby’s foot → determined foot to be small → called number → recalled messages → adopted target behaviors

For the assessment of behavioral impact, primary analysis was done by restricting the birthweight of newborns to <2,000g, and with a less restrictive birthweight cut-off of < 2,500g. Analysis of effect size was a simple difference in proportions, which was assessed for significance using chi-square test, adjusting for cluster design, using the generalized linear model with binomial distribution and log link function.

For the process variables, analysis was done by disaggregating the variables in several ways, notably by birthweights of < 2,000g and <2,500, as determined by the field staff, and foot length assessed as ≤6.9cm, as determined by the family and by the field staff. Use of the foot-length cards, calling the toll-
free number, and recall of key messages were analyzed as simple proportions with 95% confidence intervals. All analysis was performed using Stata Version 14 (StataCorp, College Station, Texas).

**Results**

All 34 clusters were randomly assigned and included in the final analysis. As illustrated in the flow diagram (fig. 5) a small number of potential individual participants were excluded or—due to missing data—were not included in the analysis of primary endpoints.

Study participants were recruited over the period July 2016 through January 2017, and each was followed up through two weeks post-delivery. As seen in Table 1, randomization of clusters achieved good balance on relevant socio-demographic and clinical characteristics.

**Table 1: Baseline characteristics**
### Table 2: Principal outcomes, care practices for low birthweight newborns (34 clusters)

|                                | Comparison arm | Intervention arm |
|--------------------------------|----------------|------------------|
| **Sex of the baby**            |                |                  |
| Male                           | 53.2%          | 52.3%            |
| Female                         | 46.8%          | 47.7%            |
| **Total**                      | **2,432**      | **2,022**        |
| **Age of mother at delivery (years)** |            |                  |
| <20                            | 25.4%          | 25.0%            |
| 20-24                          | 43.5%          | 43.2%            |
| 25-29                          | 21.7%          | 22.5%            |
| 30-34                          | 6.9%           | 6.8%             |
| ≥35                            | 2.6%           | 2.5%             |
| **Total**                      | **2,432**      | **2,142**        |
| **≥1 antenatal visit**         |                |                  |
| No                             | 13.3%          | 13.8%            |
| Yes                            | 86.4%          | 86.1%            |
| **Total**                      | **2,424**      | **2,141**        |
| **Delivered in health facility** |              |                  |
| No                             | 47.4%          | 47.9%            |
| Yes                            | 52.6%          | 52.1%            |
| **Total**                      | **2,431**      | **2,142**        |
| **Number of times pregnant, prior to this delivery** |        |                  |
| None                           | 26.5%          | 26.4%            |
| 1                              | 28.0%          | 28.4%            |
| 2-4                            | 39.6%          | 39.8%            |
| ≥5                             | 5.7%           | 5.3%             |
| **Total**                      | **2,430**      | **2,141**        |
| **Years of education – mother (years)** |          |                  |
| None                           | 63.4%          | 63.5%            |
| 1-4                            | 4.7%           | 4.3%             |
| 5-9                            | 19.1%          | 18.8%            |
| ≥10                            | 12.7%          | 13.4%            |
| **Total**                      | **2,430**      | **2,141**        |
| **Years of education – husband (years)** |          |                  |
| None                           | 40.1%          | 39.9%            |
| 1-4                            | 7.6%           | 5.8%             |
| 5-9                            | 30.8%          | 33.7%            |
| ≥10                            | 21.3%          | 20.5%            |
| **Total**                      | **2,430**      | **2,139**        |
| **Wealth terciles (by asset possession)** |        |                  |
| Low                            | 23.1%          | 25.2%            |
| Medium                         | 48.6%          | 48.1%            |
| High                           | 28.2%          | 26.7%            |
| **Total**                      | **2,432**      | **2,142**        |

*a Calculated by combining 5 asset-related variables: mobile phone, television, land, latrine, and motorcycle. All variables dichotomized and then summed; scores then stratified into 3 categories; high, medium and low.

As seen in Table 2, babies in the intervention arm who were born small (either <2,500g or <2,000g) were approximately 2.5 times more likely to receive skin-to-skin thermal care than those in the control arm. No significant difference was seen for care-seeking outside the home.

**Table 2: Principal outcomes, care practices for low birthweight newborns** (34 clusters)
The study also sought to open the black box on the expected causal mechanisms by which the intervention may produce its desired impact on care practices. As explained in the Methods section, our hypothesized causal process was as follows: received card → measured baby’s foot → determined foot to be small → called number → understood and recalled messages → adopted target behaviors. As is seen in Table 3, a high proportion of those in the intervention arm received the foot-length card and associated instructions, during a late pregnancy home visit. And, of those receiving the card, close to three quarters reported having used it to check their baby’s foot-length. However, of those checking foot-length, only a very small proportion (1.7%, 95% CI: 1.0-2.3) reported that they found the baby’s foot to be shorter than the card. This was in marked contrast to findings of earlier anthropometric assessments in this population, and is discordant with results from field-staff use of the cards (see Figure 6). Among the very small number of mothers reporting having called the toll-free number, message recall was relatively good (results not presented here), however given that such calls were made for only a very small proportion of low birthweight babies, the intervention—as designed—could not be expected to produce a discernable population-level effect, regardless how well understood and recalled the messages were.

Table 3: Process analysis (restricted to the intervention arm)

| Event                              | n     | % (95% CI)     | Total N |
|------------------------------------|-------|----------------|---------|
| Received card and instructions     | 1,953 | 96.6 (95.8-97.4)| 2,022   |
| Used card to check baby’s foot-length | 1,443 | 73.9 (71.9-75.8) | 1,953  |
| Judged baby’s foot smaller than the card | 24    | 1.7 (1.0-2.3)  | 1,443  |
| Called toll-free number & listened to messages | 62 | 3.2 (2.4-3.9) | 1,953 |
| Called number & listened to messages, among <2,500g | 22 | 4.6 (2.7-6.5) | 476 |
As illustrated in Figure 6, from data in an anthropometric study conducted in this population [3], using their measurement device and a foot-length cut-off of 6.9cm, for infants born <2,000, this criterion had a sensitivity of 87.5% (i.e. it would misclassify only 12.5% of these very small newborns as normal weight). For a weight threshold of 2,500g, the same cut-off had a specificity of 94.2% (i.e. of those with foot-length ≥2,500, only 5.8% would be misclassified as small). By contrast, in our study when mothers or other family members used the foot-length card to classify their newborns, they missed 84.5% of those that were actually <2,000g (classifying them as normal). When field staff used the card, their performance was intermediate between use of the measuring device and mother or family member classification, but they still misclassified 35.3% of those <2,000g as normal weight.

Discussion

The main outcome measures for the trial were reported skin-to-skin thermal care and seeking care for the newborn from a service provider outside the home, restricted to low birthweight newborns (using two thresholds: 2,000 g and 2,500 g). For care-seeking, the study findings do not support there being any intervention effect. For reported skin-to-skin care, however, the trial found a relative risk of 2.50 (95%CI: 2.01–3.1), with the 2,500 g threshold, and 2.48 (95%CI: 1.37–4.51) for 2,000 g. One could—naively—infer that the foot-length card screening/recorded-message intervention was effective in improving thermal care practices, notably skin-to-skin care. However, the study also documented intermediate steps in the presumed causal chain through which the effects of the intervention could be mediated. The weak link in that chain turned out to be classification accuracy using the card. Even when used by research field staff, who had the opportunity to build their skills assessing multiple newborns, use of the card poorly approximated the results obtained with the specially designed measurement device used in the earlier anthropometric study, seen in Fig. 1 [2]. In the hands of mothers and other family members, use of the card to assess foot-length—as a basis for classifying their newborns as at risk due to low birthweight—performed very poorly, misclassifying 84.5% of very small newborns (<2,000 g), and fully 95% of those < 2,500 g, as normal. So any enhanced adoption, for these small babies, of better thermal care practices cannot be attributed to parents identifying their baby as small using the card, calling the number, hearing the messages, understanding, and applying them.

The study used a randomized control trial methodology, sample size was adequate for the outcomes assessed, and good balance was achieved between treatment and control arms. One constraint was that it was important to interfere as little as possible with the parent study on which our trial was superimposed, minimizing additional engagement with study participants or measurement. This limited us to potentially relevant outcomes that were already being measured in the parent trial. Had the study been conducted as a stand-alone trial, a fuller range of outcome measures better approximating the objective of this intervention strategy could have been used.

Formative work had been done in a rural Nepal setting to validate that the phone script used was well understood by pregnant women and women who had recently given birth. However, in retrospect, an unwarranted assumption was made that measurement and classification by mothers or other family
members, using a card measuring 6.9 cm, would closely approximate results obtained by field research staff with a purpose-built device as used in the earlier anthropometric study. This was not tested in advance of the trial.

The findings appear paradoxical, on the face of it. On the one hand, among small newborns in the intervention arm, skin-to-skin thermal care was practiced considerably more than in the control arm. But, based on the process measures, it is clear that this cannot be attributed to correct classification, calling the number, hearing the messages and applying them. So what can account for the apparent impact of the intervention?

At the time of the first postnatal visit, of those who had received the card, only 3.2% (62/ 1,953) reported having called the toll-free number and heard the recorded messages. It is possible that over the following days, more of parents called the number, regardless of what they found using the foot-length card, and this—conceivably—could have influenced thermal care practices. However, any calls that may have occurred subsequent to the first postnatal home visit were not assessed in this study so we are not in a position to further explore this hypothesis. An alternative explanation is that some aspect of the interaction between the field research staff and study participants, either during the late pregnancy visit or at the time of the initial contact after birth, could have sensitized mothers somehow to the importance of skin-to-skin care for very small newborns, regardless how they classified them using the card. Study co-investigators who were involved in field supervision have reviewed the procedures used, but have failed to identify a plausible explanation for how this could have occurred, based on how field contacts were implemented. Of course, even with confidence intervals around the relative risk point estimate that do not overlap 1.0, it certainly remains possible that such a result could have arisen through pure chance. In our judgement, this is the most credible explanation for the study finding of an apparent intervention effect.

As noted in the Introduction, newborns of low birthweight account for the majority of neonatal deaths [3], many of which could be prevented with good attention to thermal care [4], optimal newborn breastfeeding practices [5], and prompt medical attention for danger signs. In settings where many births still occur at home, not attended by professional health workers, there is a compelling program logic to trying to find ways of identifying these small newborns and ensuring the needed care practices. In such a context, simple tools that could allow visiting community health workers, or mothers themselves, to identify higher risk babies, have real appeal. Salter spring scales, made specifically for weighing infants, are inexpensive and comparatively simple to use. However, for accurate results they do require skills and proper procedure; earlier efforts in Nepal to have FCHVs make postnatal home visits, which included weighing newborns, faltered, both due to low early home visitation rates and difficulties the FCHVs had using the scales [9].

Newborn anthropometric studies have been done in a number of settings, assessing correlation between birthweight and a variety of other measures, including foot-length, and have generally shown reasonably good test characteristics for foot-length [2, 12, 15, 16, 17, 18]. Although most of these studies did not find
foot-length to be the best proxy for birthweight, foot-length measurement has seemed more practical for field use than other measures that generally performed better (e.g. chest or head circumference).

Two groups have tested strategies involving having CHWs assess foot-length, during home visits. As noted in the Introduction, in a field study in Tanzania [15], during postnatal home visits, CHWs assessed foot-length to classify babies and, if small, gave counseling messages on needed care. In the study, CHWs first classified cases using the card, then measured the baby's foot using a hard ruler. Subsequently, the field researcher independently assessed foot-length using the card, and then measured using the ruler. Inter-rater reliability was assessed, comparing length measures made by the CHWs and the field researcher. Moderate reliability was found, with a kappa of 0.53. On average, CHWs assessed foot length at 0.3 cm shorter than measures by the field researcher, thus overestimating the number of small babies (opposite to the problem found in the current study). In this Tanzanian study, length measures were not assessed against weight (a more appropriate gold standard).

Another study, conducted in Bangladesh [19], assessed performance of CHW assessments done during early postnatal home visits, which included foot length measures with a hard ruler, using the same procedures as in the Tanzanian study. For the Bangladesh study, however, the primary interest was correlation with gestational age, as determined by early pregnancy ultrasound. Foot-length, as measured by CHWs, performed poorly in this regard.

With such findings, along with the results of the current study, appealing though the idea may be, having family members or CHWs assess risk associated with low birth weight using foot-length as a proxy does not appear to be a promising strategy. The challenge remains of how best to identify babies at risk, and to mitigate such risks through targeted efforts to ensure optimal care practices. In principle, for institutional births, mothers of all babies born at a weight that puts them at higher risk need to be given good counseling, before discharge home, on key practices, notably: thermal care, breastfeeding, and danger sign recognition and associated prompt care-seeking. For populations in which home births, not attended by professional health workers, remain common, other strategies are needed. Counseling during late pregnancy antenatal contacts should certainly include essential newborn care messages, emphasizing particular risks faced by smaller-than-average newborns.

**Conclusion**

On one of the two primary outcomes in this community-based cluster randomized trial, skin-to-skin thermal care, participants in the study arm were 2½ times more likely to report the practice than those in the control arm. However, process measures along the hypothesized causal chain document that the intervention did not perform well: classification of newborns as small using a card 6.9 cm long (a cut-off determined from an earlier anthropometric study in this setting) performed poorly as a proxy for birth weight. As a result only a very small proportion of those who could have benefited from doing so followed up to make a call to receive information on care of these higher-risk newborns (including on skin-to-skin care). So, very few of those with small newborns in the intervention arm were actually exposed to the
health messages that constituted the core of the intervention. We cannot rule out the possibility that some aspect of conduct of the study within the intervention arm could have influenced women in the intervention arm to adopt this thermal care practice. But—on review of study field procedures—it seems more plausible to us that this result is attributable to chance.

Making an unbiased, reasonably accurate classification of foot-length using a card, as we attempted to have household members do in our study, is not as straightforward as we expected. There is still a need, in Nepal and other similar settings, to reach women giving birth at home, to inform and empower them to take action to reduce risk for babies born very small, but a different approach will be needed.

**Abbreviations**

CHW – Community Health Worker

FCHV – Female Community Health Volunteer

VDC – Village Development Committee

**Declarations**

**Ethics approval and consent to participate**

Participants were given verbal instructions and an information sheet on the study during late pregnancy visits conducted by the trial field staff and then—if they agreed—they signed the informed consent form (or indicated consent with a mark). These documents were in the local languages (Maithili and Nepali) and were read aloud to women unable to read. Note that no participants were younger than 16 years old.

Ethics approval for the trial was received from institutional review boards at the Johns Hopkins Bloomberg School of Public Health and the Institute of Medicine, Tribhuvan University in 2016. In both cases, this was submitted as an amendment to the broader protocol for the parent trial.

**Consent for publication**

Not applicable

**Availability of data and materials**

The datasets used and/or analysed during the current study are available from the corresponding author on reasonable request.

**Competing interests**

The authors declare that they have no competing interests

**Funding**
This study was funded by Save the Children/ USA, under the Saving Newborn Lives project. At the time the trial was conducted the authors (SH, BR, DJ, BB) were employed by Save the Children.

**Authors' contributions**

Study design was led by SH and LM with contributions by BR, DJ, BB, SK; field implementation of the trial was led by SK, with support from BR; analysis was done by DJ, with support from LM; the manuscript was written primarily by SH, with input from BR, DJ, BB, SK and LLM; all authors reviewed and approved the final manuscript.

**Acknowledgements**

Field staff of the Nepal Nutrition Intervention Project/ Johns Hopkins University were responsible for field implementation. Dr. Anjana KC and Elaine Scudder (Save the Children) provided management support for the study.

**References**

1. Black RE. Global Prevalence of Small for Gestational Age Births. Nestlé Nutr Inst Workshop Ser. 2015;81:1–7. doi: 10.1159/000365790

2. Mullany LC, Darmstadt GL, Khatry SK, Leclerq SC, Tielsch JM. Relationship between the surrogate anthropometric measures, foot length and chest circumference and birth weight among newborns of Sarlahi, Nepal. Eur J Clin Nutr. 2007;61:40-6.

3. Lawn JE, Cousens S, Zupan J, et al. 4 million neonatal deaths: when? Where? Why? Lancet. 2005;365:891-900.

4. Mullany LC, Katz J, Khatry S, LeClerq SC, Darmstadt GL, Tielsch JM. Risk of Mortality Associated With Neonatal Hypothermia in Southern Nepal. Arch Pediatr Adolesc Med. 2010; 164:650–656. doi: 10.1001/archpediatrics.2010.103 PMID: 20603466

5. Mullany LC, Katz J, Li Y, Khatry SK, LeClerq SC, Darmstadt GL, Tielsch JM. Breast-feeding patterns, time to initiation, and mortality risk among newborns in Southern Nepal. J Nutrition. 2008;138:599–603.

6. Chang KT, Mullany LC, Khatry SK, LeClerq SC, Munos MK, Katz J. Validation of maternal reports for low birthweight and preterm birth indicators in rural Nepal. J Glob Health. 2018;8:010604. doi: 10.7189/jogh.08.010604.

7. Ministry of Health, New ERA, ICF. Nepal Demographic and Health Survey 2016. Kathmandu: Ministry of Health Nepal; 2017.

8. Penfold S, Sharma G, Joshi N, Adhikari M, Shapiro M. Chlorhexidine Coverage and Compliance Survey. Kathmandu, Nepal: CAMRIS International; 2017.

9. McPherson R. Assessment of the Community-Based Newborn Care Package in Nepal. Washington, DC: MCHIP, Jhpiego; 2013.
10. Schumann DA, Mosley WH. The household production of health. Introduction. Soc Sci Med. 1994;38:201-4.

11. Marchant T, Penfold S, Mkumbo E, Shamba D, Jaribu J, Manzi F, Schellenberg J. The reliability of a newborn foot length measurement tool used by community volunteers to identify low birth weight or premature babies born at home in southern Tanzania. BMC Public Health. 2014;14:859. doi:10.1186/1471-2458-14-859.

12. Pratinidhi AK, Bagade AC, Kakade SV, et al. Action-oriented colour-coded foot length calliper for primary healthcare workers as a proxy for birth weight & gestational period. Indian J Med Res. 2017;145:347-352. doi: 10.4103/ijmr.IJMR_36_14.

13. KC A, Nelin V, Vitrakoti R, Aryal S, Målvqvist M. Validation of the foot length measure as an alternative tool to identify low birth weight and preterm babies in a low-resource setting like Nepal: a cross-sectional study. BMC Pediatr. 2015;15:43. doi: 10.1186/s12887-015-0361-4.

14. Mullany LC (trial posted 2010). Impact of Sunflower Seed Oil Massage on Neonatal Mortality and Morbidity in Nepal (NOMS). (ClinicalTrials.gov: NCT01177111) Recruitment closed Jan 2017; results not yet published.

15. Marchant T, Jaribu J, Penfold S, Tanner M, Armstrong Schellenberg J. Measuring newborn foot length to identify small babies in need of extra care: a cross sectional hospital based study with community follow-up in Tanzania. BMC Public Health. 2010;10:624. doi: 10.1186/1471-2458-10-624.

16. Nabiwemba EL, Garimoi C, Kolsteren P. Determining an anthropometric surrogate measure for identifying low birth weight babies in Uganda: a hospital-based cross sectional study. BMC Pediatr. 2013;13:54. doi: 10.1186/1471-2431-13-54.

17. Rustagi N, Prasuna JG, Taneja DK. Anthropometric surrogates for screening of low birth weight newborns: a community-based study. Asia Pac J Public Health. 2012;24:343-51. doi: 10.1177/1010539510384717.

18. Thi HN, Khanh DK, Thu Hle T, Thomas EG, Lee KJ, Russell FM. Foot Length, Chest Circumference, and Mid Upper Arm Circumference Are Good Predictors of Low Birth Weight and Prematurity in Ethnic Minority Newborns in Vietnam: A Hospital-Based Observational Study. PLoS One. 2015;10:e0142420. doi:10.1371/journal.pone.0142420

19. Lee AC, Mullany LC, Ladhani K, Uddin J, Mitra D, Ahmed P, Christian P, Labrique A, DasGupta SK, Lokken RP, Quaiyum M, Baqui AH; Projahnmo Study Group. Validity of Newborn Clinical Assessment to Determine Gestational Age in Bangladesh. Pediatrics. 2016;138(1). pii: e20153303. doi:10.1542/peds.2015-3303.

**Figures**
Figure 1

Device used in earlier study [2]
Figure 2
Counseling/ measurement card (from Marchant 2014 [11])

Figure 3
Measurement caliper (from Pratinidhi 2017 [12])
The Foot-Length Card. Note: Text of instructions given (in English): Concerning this card ... For babies born very small there can be more danger than for normal sized babies. This card is used to help to tell if your baby is very small. This is how to use it ... After your baby is born (whether at home or health facility), you should check to see if your baby's foot is longer than the card. Placing the baby on a flat surface, put the card against the sole of the baby's foot, resting the bottom of the card on the surface the baby is lying on. Then you can check whether or not the baby's toes extend beyond the end of the card. If the baby's foot is NOT longer than the card or even if the foot is longer but the baby still seems to you to be very small, call one of the numbers on the card. The call is free. When you call, you will be able to listen to a message on special care needed for the baby to help protect him or her. You can listen as many times as you like.
Figure 5

Participant flow diagram (34 clusters: 17 each randomized to intervention vs. control)
Figure 6

Classification Accuracy for Low Birthweight (using a 6.9cm cut-off)