

Chapter 2

Methodology

This chapter describes the methodology that has been adopted to address the evaluation questions outlined in the previous chapter. We have stated in the previous chapter that the broad objectives of our evaluation exercise are of two-fold: On the one hand, we aim to evaluate the training programme conducted so far; on the other hand, we intend to examine the need, design, effectiveness and efficiency of the training programme by using a case-controlled semi-randomised design.

The sections of this chapter are organized as follows: Section 2.1 clearly delineates the basic approach followed in this evaluation exercise. Section 2.2 describes design of the baseline survey. Section 2.3 describes how we go about evaluating the need, design, effectiveness and efficiency of the study. Sampling techniques and sample sizes for both the designs are described in the respective sections. While summarizing the main points outlined in sections 2.1 and 2.2, Section 2.4 points out the limitation of the evaluation design which one should be aware of w

2.1 Basic Approach to Evaluation

Our approach for evaluation is an independent one in which we - the evaluators- take the primary responsibility of developing the evaluation design, collect and analyse the data and disseminate the findings. However, all through our evaluation process we have given due considerations to accommodate any particular issue that has been raised by either the implementing agency (i.e. Liver Foundation, West Bengal) or the funding agency (i.e. BMS Foundation) or both. Further, we have also attempted to incorporate the perspectives of all stakeholders who have been directly or indirectly associated with the training programme or its outcomes. Since the direct and indirect target groups of the training programme have been RHCPs and the population served by them, the evaluation also consider them as direct and indirect target groups respectively

The evaluation design splits the target population into two sub-groups keeping various constraints in mind: (a) the experiment group – the population which is covered by the training programme (i.e. intervention); and (b) the control group – the population which is not covered by the training programme. Given the complexity, sensitivity and other contextual constraints, we resort to semi-randomised or quasi-randomised design. However, sufficient attention has to be given to avoid selection and other biases in estimation of programme effects that might result from semi-randomised design. Adequate focus has also been given to process evaluation (i.e. assessing and monitoring the training programme process).

The RHCP training programme was started in 2007-08 in 2 blocks of Birbhum district in West Bengal. Till 2010-11 a total number of 815 RHCPs in 16 blocks spread in two states (namely, West Bengal and Jharkhand) were selected for the training and out of them 689 RHCPs completed the training programme. Till 2010-11, the training covered 13 blocks in West Bengal and 3 blocks in Jharkhand (see Table 2.1).

Table 2.1: Block-wise number of RHCPs trained in West Bengal and Jharkhand (2007-2011)

States	Districts	Blocks	No of RHCPs trained
West Bengal	Birbhum	Rajnagar	42
		Suri I	43
		Khoyrasole	59
		Md Bazar	63
		Dubrajpur	25
		Mayureswar I	35
		Mayureswar II	25
	Nadia	Nakashipara	46
		Chapra	35
	Purulia	Arsa	40
		Purulia 2	35
	South 24 Parganas	Pathar Pratima	48
	Murshidabad	Nabagram	50
Jharkhand	Dumka	Moslia	29
		Raniswar	38
		Shikaripara	31

Source: Information provided by Liver Foundation (Kolkata and Suri offices)

Out of all the blocks which were covered by the RHCP training programme, we selected 9 blocks considering their diversity or logistic aspect to conduct surveys. More than 200 RHCPs and more than 1000 households were selected in different rounds of the survey. While making two groups of RHCPs comparable we adopted score matching method. Score given to each RHCP (independent of selecting for the interview) depended on three basic information: (a) educational qualification (measured by years of schooling) and (b) years of experience as an RHCP. A composite index was then calculated combining education and experience To ensure an unbiased representation of the trained RHCPs from block, at least one-third of them were selected spreading over the whole range of composite index values within that block.

2.2 Evaluation Design for the baseline Survey

The sample design for our baseline survey selected three blocks/ part of blocks with the aim of stratifying the Gram Panchayats (the lowest layer of local government) within the blocks into experimental and control groups. Selection of the blocks was guided by a few practical considerations. Before the evaluation was started, the training programme covered almost half of the Gram Panchayats of Dubrajpur block. Dubrajpur is a block which has close proximity to Dubrajpur Municipality area – an urban inhabitation. Dubrajpur block is also better connected with district headquarter Suri by state highway. At the time of selecting few Gram Panchayats from Dubrajpur block, it was decided by the training organization that in the next round of the training rest of Dubrajpur block would be covered. Therefore, we included the remaining 5 Gram Panchayats of Dubrajpur block as a part of the experimental group which were left out of training. The second block that we selected was Mayureswar II. Mayureswar is predominantly a rural area with no municipality. There are 7 Gram Panchayats in Mayureswar II. We divided all Gram Panchayats in Mayureswar II block into two comparable groups: one group consists of 4 Gram Panchayats and the other group consists of three.

In selecting comparable Gram Panchayat we first created a composite index for each Gram Panchayat based on three sets of information: share of marginalized/vulnerable group in the population (33.3% weight), population served per Sub-Centre/PHC (16.7%

weight), persons served per RHCP (16.7% weight) and the distance of the GP from the nearest town (33.3% weight). Groups of comparable Gram Panchayats were created by allocating Gram Panchayats with similar combined score with both the groups.

The third block we selected was Sainthia. Sainthia was selected in order to include Gram Panchayat having characters of the Gram Panchayats belonging to Dubrajpur block. Like Dubrajpur, Sainthia also has an adjacent Municipality area. Given the fact that Panchayats in Dubrajpur area were already selected, we tried to select a set of Gram Panchayats from Sainthia which were more-or-less comparable to Dubrajpur. In so doing we considered all the Gram Panchayats belonging to Sainthia block and then we computed combined score for each of the Gram Panchayat. Similar scores were generated for Gram Panchayats belonging to Dubrajpur block. We selected only those Gram Panchayats in Sainthia block which were similar to the Panchayats in Dubrajpur in terms of combined score. The list of experimental and control group Gram Panchayats in Dubrajpur, Mayureswar and Sainthia blocks are presented in Table 2.2.

Table 2.2: List of experimental and controlled group Gram Panchayats in Dubrajpur, Mayureswar II and Sainthia blocks.

Experimental Group Gram Panchayats	Controlled Group Gram Panchayats
Balijuri (Block: Dubrajpur)	Ahmadpur (Block: Sainthia)
Joshpur (Block: Dubrajpur)	Amarpur (Block: Sainthia)
Loba (Block: Dubrajpur)	Banagram (Block: Sainthia)
Paruliya (Block: Dubrajpur)	Bromorkol (Block: Sainthia)
Puduma (Block: Dubrajpur)	Hatora (Block: Sainthia)
Dekha (Block: Mayureswar II)	Sangra (Block: Sainthia)
Koleswar (Block: Mayureswar II)	Daspalsa (Block: Sainthia)
Kundula (Block: Mayureswar II)	Mayureswar (Block: Sainthia)
Ulkunda (Block: Mayureswar II)	Satpalsa (Block: Sainthia)

Source: Sample design for evaluation.

Once the experimental and controlled group Gram Panchayats were selected, a list of RHCPs with education 10 years of schooling or more and willing to participate in the survey (and later training for the experimental group) was prepared. The baseline survey covered almost all the RHCPs who were included in the list. Although the baseline survey interviewed 110 RHCPs (49 belonging to experimental group and 61 belonging to controlled group), we excluded 6 RHCPs from the analysis as they were found practicing

mostly Homeopathy medicines. Our final analysis of the baseline data was based on interview of 104 RHCPs (47 belonging to experimental group and 57 belonging to the control group).

We relied on the interviewed RHCPs to select their users (i.e. the households). However, adequate precautions were taken to avoid bias as far as possible. The selected RHCPs were requested to provide a list of their users under different stratifications: Hindu, Muslims, Scheduled Castes, Scheduled Tribes, Rich, Poor, Households with Children, Household with Pregnant Women etc. Households were randomly selected from the stratified lists which were prepared in consultation with the RHCPs. Generally 7-10 users/ households were chosen for each RHCP. We tried to interview as many ANMs and GP members as possible. Whereas it was relatively easier to convince a GP member for an interview, ANMs were not easily available for finding time for the interview.

2. Design for Evaluating RHCPs Training Carried Out So Far

One of the requirements of this evaluation exercise, especially from the funding agency's point of view was to assess the training programme carried out so far. Since there was no systematic baseline survey done before starting the training programme, we decided to select a good proportion of 'already trained' RHCPs who were spread in different blocks and were trained by different group of trainers. Since it was not feasible to know about the knowledge, attitude and practice of the RHCPs before they joined the training programme, we resorted to various retrospective questions to have some idea about the situation before the training programme.

In order to assess the 'training carried out so far', we focused on the training related experience of the trained RHCPs and tried to base our analysis from a larger sample in order to capture as diverse views as possible. As per the data provided by the Liver Foundation, 689 RHCPs have successfully completed training programme in different blocks of West Bengal and Jharkhand till 2010-11 and our analysis is based on a sample of 128 RHCPs which is roughly 18.6% of total number of RHCPs that were provided training by the Liver Foundation till 2010-11. Since our sample of 128 RHCPs come

from three different districts and were trained by different group of trainers in each district, it may be appropriate to present the some of the indicators separately for the districts so that trainer-group specific effects on the RHCPs can be partially analysed.

2.4 Evaluation based on Semi-experimental Design

This part of the analysis compares an experimental group and a control group before and after the training programme (i.e. intervention) and tries to bring out their differences in terms of their knowledge, attitude and practices. Although our baseline survey interviewed 104 RHCPs practicing allopathic system, only 85 of them were available for the endpoint survey. The split of 85 RHCPs between experimental and control group became 27 for the experimental group and 58 for the control group. In other words, the size of our experimental group of RHCPs is little less than half of the control group.

Since it is always feasible on the part of the training organization to artificially improve the performance of the experimental group RHCPs only by emphasizing on those topics on which questions were asked in the baseline survey, we included three more stakeholders in our survey design – users of the RHCPs, ANM (government health workers) and Gram Panchayat Members. The comparison of pre-training and post-training experimental and controlled user groups are based on information from 633 households in both rounds of survey (viz. pre-training and post-training). Out of 633 households, which finally remained part of the baseline survey, 220 households together formed the experimental group, while 413 households together formed the controlled group. For the post-training household surveys, 633 households remained part of the final analysis for comparison of which 216 households formed the experimental group and 417 households formed the experimental group. RHCPs mostly provide curative health care whereas ANMs functioning in the same areas offer mostly preventive care to the rural population. As it became difficult to get good number of ANM for the interview, our pre-training and post-training surveys could interview only 48 and 47 ANMs respectively. The split between experimental and controlled group ANMs for the baseline (pre-training) survey was 28 and 20, it was 27 and 20 for the endpoint (post-training) survey. It was relatively easier to get more number of Gram Panchayat members for the

interview. As a result, we could interview as large as 188 GP members both during pre-training and post-training surveys. Out of 188 GP members interviewed for the baseline survey, 100 GP members formed the experimental group, while remaining 88 GP members formed the control group. Out of the 188 GP members interviewed for the endpoint survey, 96 GP members together formed the experimental group, while remaining 92 GP members together formed the control group.

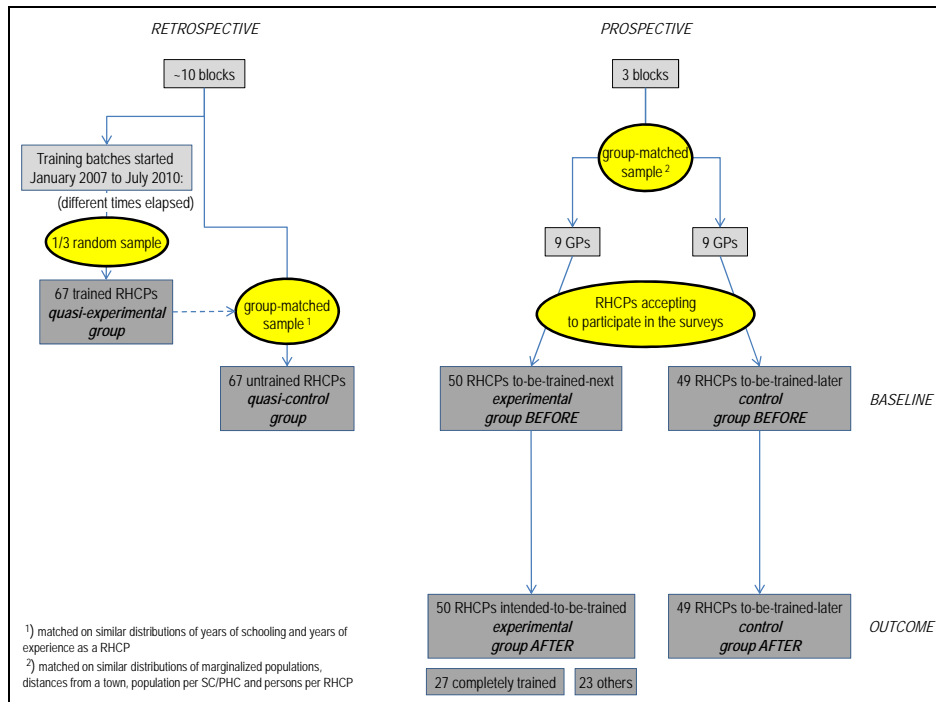
APPENDIX

A Note on Evaluation Methodology and Sample Design¹

Prepared by Dr Rolf Heinmuller (University of Montreal)

Most of the **RHCP training batches** were completed or engaged at the time the evaluation research was planned. Two out of some 17 blocks had to be excluded from the survey, because political unrest might have perturbed the data collection: Arsha and Khoirasole. In the participating blocks, 67 *trained* were randomly sampled and group-matched with 67 *untrained* RHCPs such that both groups had similar distributions of years of schooling and of service as RHCP. Obviously, the already completed or engaged batches could only be evaluated **retrospectively**; the evaluators could not intervene to increase the comparability of catchment areas nor of trained and untrained RHCPs (which were termed **quasi-experimental** and **quasi-control** groups in the interim report of the evaluation study; see figure A 2.1 below, left half).

Figure A 2.1: Evaluation design

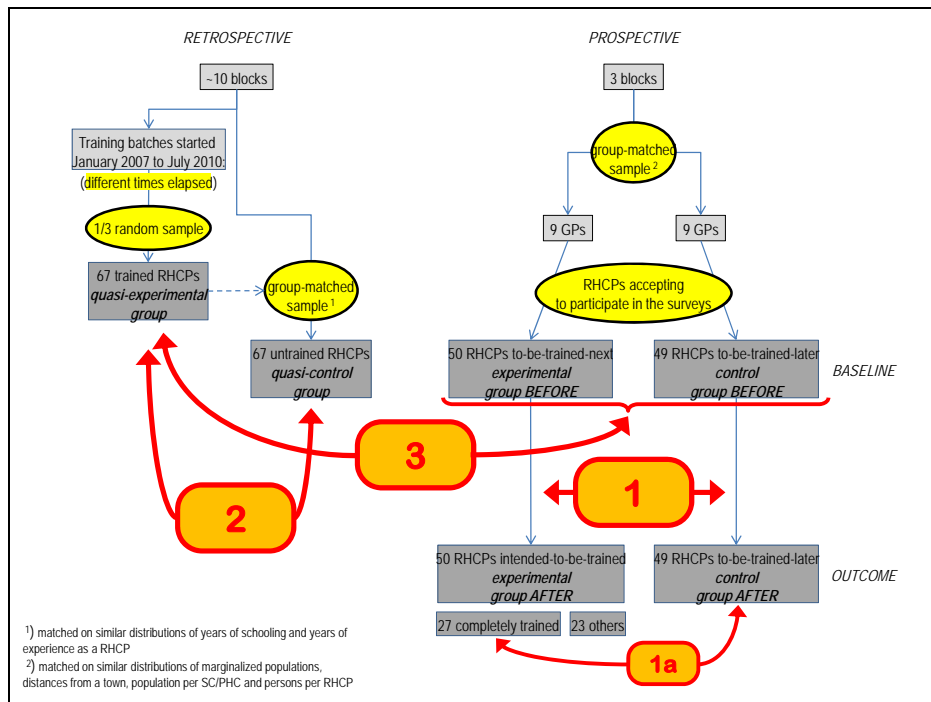


¹ This note is based on the sample design that was adopted for preparing the interim report. The note also guides for the scientific papers that will be based on the survey data.

By contrast, the next two batches had not yet been selected when the evaluation study was planned. The evaluators had the chance to control their composition. They contributed to their selection by increasing the comparability of trainees and areas between the two batches. The 18 gram panchayats to be included in the two batches were divided so as to *prospectively* equilibrate certain characteristics of the areas (see figure A2.1, right half). The two groups were to have similar distributions of the following four parameters: share of marginalized populations, distance to the nearest town, population served by Sub-Centre or PHC, and persons served by RHCP (the former two criteria were weighed 33.3 % each, the latter two, 16.7 % each); insofar the two groups are randomized. One batch (the *experimental* group) to was surveyed *before and after* training, while the training of the other batch (*control* group) was postponed until after the second survey, in order to maintain a comparable, untrained control group.

The prospective groups enable the most controlled *analysis no. 1* (see figure A2.2). The *experimental* and *control* RHCP groups are reasonably comparable with one another, because both contain RHCPs that could turn out eligible or not for training, and their contexts are equilibrated. This analysis uses the maximum level of randomized allocation that was possible at the time the evaluation was planned. It estimates the effect of training in those RHCPs to whom it was offered (whether they complete it or not) and corresponds to what is called *intention-to-treat* analysis in randomized controlled trials. By contrast, *analysis no. 1a* compares only the RHCPs who have *actually completed* the training with the controls and estimates the effect in 100%-compliant trainees. The groups of this comparison are *not* randomized, their comparison estimates a kind of best-case effect. — Analyses *1* and *1a* imply the longitudinal, before-after comparison (not shown in figure A2.2) of experimental and control groups, which controls for history and maturation biases.

Figure A.2.2: Comparisons

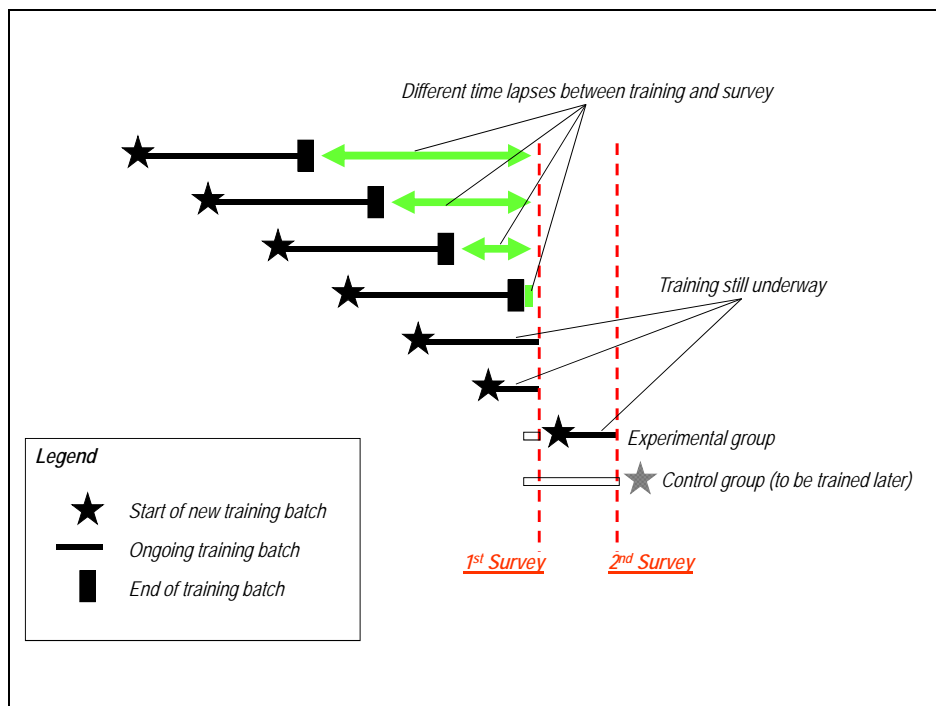


Analysis no. 2 designates the quasi-experimental part of the design. The *quasi-control* RHCPs may obviously differ from the *quasi-experimental* RHCPs in many respects, notably due to the admission criteria (see section 2); this aspect requires careful consideration during analysis and interpretation. They represent a more knowledgeable part of the hitherto excluded RHCPs. Furthermore, in this analysis it will be important to control the results of different batches of quasi-experimental RHCPs for the different times elapsed since the beginning of their training courses (see figure 3, below). Analysis 2 might enable obtaining an impression of the *longtime effect* of the training and possibly increasing or fading shape, if any (which is of course unavailable from analysis 1) — this possibility may turn out to be limited or facilitated by other features of the data.

Analysis 3 opposes the quasi-experimental group to the *pooled* two prospective groups as they were observed at baseline (i.e. before the experimental group started training). These groups are more comparable than the groups opposed in analysis 2, because both experimental and control RHCPs were all yet untrained but mostly *eligible* for training (other than the quasi-control group). Their comparison enables estimating a similar effect as does analysis 1. However, the contexts of this comparison have not been equilibrated and the effect may be modified by the times elapsed for the less recent batches. In turn, this analysis has more power at the level of individual RHCPs

than analysis 1, because there are many more of them ($n = 134 + 99$). If, instead, quasi-experimental and the 27 compliant members of the experimental group (as observed at the 2nd survey, see figures 1 and 2, bottom center) are pooled and opposed to the control group ($n = 161 + 49$, **analysis 3a** not shown in figure 2) and if the time elapsed can be controlled for, then one may be able to relativize the effect obtained in analysis 1a with respect to *variation between batches*, an independently relevant additional information about the robustness of the training effect.

Figure A 2.3: Different time relationships between training and survey



Apart from trained and untrained RHCPs, the evaluation study considered patients, public health workers and community representatives (these are not shown in figures A2.1 and A2.2).

Patients were selected for the evaluation study by asking every participating RHCP to list patients by whom they had most recently been consulted. This dependent sampling procedure is much more efficient than trying to capture recent RHCP users from a representative household survey. Moreover, it may capture the more positive encounters (as perceived by the RHCPs), to measure user perception and satisfaction. In turn, the sampled patients may exclude some “worst cases”; they also cannot represent all RHCP users nor the health problems presented to RHCPs

nor the health seeking behavior in general. However, respective information can be obtained from existing surveys that represent the general population.

In every block from which RHCPs were trained (or which were scheduled for the recruitment of new trainees), *ANM* or *auxiliary nurse/midwife workers* and GP or *gram panchayats representatives* were interviewed. Again, these could only be interviewed retrospectively for the training batches already completed or engaged, whereas they could in principle be contacted before *and* after training for the experimental group (and before only, for the control group). However, it turned out difficult to contact *the same* ANM workers and GP representatives during 2nd survey as those seen during the 1st survey, which may attenuate the before-after comparability of the data they contributed. The comparison of ANM and GP responses between blocks should also account for the varying proportions of RHCPs admitted across batches; analysis will take care of this aspect (see section 4.c, below).

All these differences (and “incomparabilities”) are facts of life that a “*real-world*” evaluation must deal with, rather than trying to establish unrealistic, randomized settings (Bamberger & al., 2006). In this perspective, the data of the “*quasi*” groups add relevant variation and enrich the qualitative information, although they need more careful control during analysis and more caution during the interpretation of results than does the randomized analysis no. 1. The dozen “quasi” batches may also compensate for the lack of statistical power of the only two “prospective” batches.



(above) An abandoned Primary Health Centre (PHC). (Below) A functioning PHC

