

Randomization Strategy for the Eritrea Malaria IE Survey

Random Village Selection

Dr Joseph Keating randomly selected treatment and control villages in the following way:

- He received the list of eligible villages from field coordinator Abdoulaye Sy, who obtained it from the Eritrean Ministry of Health (MoH) and National Malaria Control Program (NMCP).
- STATA was used to randomize the villages into treatment and control groups (58 in each group) using the “sample” command;
- Simple random sampling was used to divide selected villages into two groups;
- Afeworki Araya, from NMCP, verified the distance between villages and replacements were made where security, border, or malaria situation was deemed inappropriate; the closest village was used as replacement.
- At the time of treatment, some villages had to be replaced, since the list initially obtained from MoH was not up to date; the closest village was used as replacement, in case the originally chosen one had moved and could not be found or reached.

Random Household Selection

Field coordinator Andrea Locatelli prepared the following randomization strategy for the selection of the households (hh) to interview in each village.

Household Enumeration

- Information was obtained on the total number of recorded hh living in the village;
- If such number exceeded the agreed upon threshold of 200, then the village was segmented to obtain segments with no more than 200 hh, allowing for at most 250 hh. E.g. if a village reportedly had 1,126 hh, then it was split into 6 segments;
- One segment was randomly chosen by the team of enumerators e.g. flipping a coin;
- All hh in the segment were listed and each was assigned a number;
- Ideally the segment should have no more than 200 hh, but up to 250 hh were allowed;
- If more than 250 hh were found in the segment, then it was split again into two sub-segments to keep the number of hh within the agreed upon limits.

Household Selection

Background

Given delays in obtaining his travel permit, the field coordinator could not leave Asmara at the beginning of the survey period, so he worked from the World Bank office in the capital. Also, it was impossible for the field teams to have access to computers in the villages to randomize there. Finally, mobile phone coverage was absent in many villages included in our study, which made it often impossible to communicate between Asmara and the field teams.

To overcome these obstacles, and in an attempt to mimic as much as possible random sampling for each village, the following strategy was devised.

Randomization Strategy

- Randomization tables were prepared in Asmara and distributed to the teams before their departure for the field;
- Microsoft Excel 2007 was used to prepare such tables;
- Using the function “=RANDBETWEEN(1,2)”, either Table 1 or Table 2 was assigned to each of the 115 villages in the study. Please refer to the attached file “TABLE CHOICE.pdf”;
- Table 1 and Table 2 provide random numbers between 1 and 20, 30, 40... 250 in each column. Please refer to the attached files “TABLE 1 RANDOM.pdf” and “TABLE 2 RANDOM.pdf”;
- Table 1 and Table 2 were prepared in the following way, following directions from Dr Pedro Carneiro:
 - For a village with a number N of hh:
 - A column was created in Excel with numbers between 1 and N;
 - Another column was created with N random numbers between 0 and 1, using the function “=RAND()”;
 - These random numbers were ranked, using the function “=RANK()”;
 - The hh corresponding to the 15 lowest random numbers were selected for the interviews.
 - Because it was not feasible to do this for each number N=1,2,...,250, each column was prepared for intervals from 11 to 20, 21 to 30, 31 to 40,..., 241 to 250;
 - To make sure each column contained at least 15 different random numbers for each segment size not exceeding 250, 24 random numbers were provided for each interval.
- For the sake of clarity, Table 1 provides instructions for use and an example, as follows:
 - Use this table to select the hh to be interviewed in every village, depending on the number of hh that live there;
 - We will admit up to 250 hh per cluster given that segmentation will be rough, and so accidentally more than 200 hh may be present after segmentation;
 - E.g. If a village has 142 hh refer to column for villages with 141-150 hh to select the 15 hh you need;
 - Please notice that the 13th number is 144 which is > 142. So you will skip this number and you will use the 16th number provided, namely 75;
 - This is why you are provided with 24 random numbers in every column, instead of just 15. You will never need more than 24.