Analysis plan of quantitative data

We describe the analysis plan for the two main outcomes: ulcer prevalence and ulcer area. Analysis for the other outcomes will follow the same model-based approach, with appropriate change to the distribution and link function if required. We will use a Bayesian generalised linear mixed model methodology for analysis to account for within-person correlation over time and correlation of outcomes within villages. Specifically, our baseline specification for person I in the village j at time t the mean function is :

$$\mu_{ijt} = \beta_0 + \beta_1 x_{ijt} + \gamma D_{jt} + \alpha_i + \alpha_j + f(t)$$

where x_{ijt} is a vector of individual level covariates (age, gender, duration of impairments and level of disability (EHF)). D_{jt} is an indicator for whether the village has received the intervention at time t, $\alpha_i \sim N(0, \sigma_1^2)$ and $\alpha_j \sim N(0, \sigma_2^2)$ are individual and village level random effects, respectively, and $f(t)$ is a function of time. We will compare different models with different specifications for time and also allow for more complex random effects structures, such as temporal decay in correlation at individual level and village-time period effects. Models will be compared using posterior predictive model checks and the leave-one-out cross-validation criterion.

We will specify weakly-informative prior distributions for the model parameters. There is little available strong evidence to support informative priors in this setting; weakly informative priors limit the plausible range of the model parameters to facilitate computation, but provide little information within this range. In particular, we will use half t_2 priors on the standard deviation terms which allow for relatively large values, which may be possible for the cohort terms.

To complete the models specifications, for the prevalence outcome, where y_{ijt}

is equal to one if the patient has an ulcer at that time and zero otherwise:

$$y_{ijt} \sim Bernoulli\left(\Lambda\left(\mu_{ijt}\right)\right)$$

where Λ is the inverse logit function. For ulcer area, we condition on having an ulcer an exclude those who do not so the outcome, y_{2ijt} is strictly positive:

$$y_{2ijt} \sim N\left(\mu_{ijt}, \tau^2\right)$$

Sample size

We will include all Primary Health Centres (PHCs) from three selected blocks for intervention (Nawagarh, Champa and Akaltara). Each block comprises 4 to 6 intervention areas defined as "health posts - PHCs" and each of these areas has between 5 and 10 people affected by leprosy who are at risk of an ulcer and another 5 to 8 patients with impairments in eye, hands with or without impairment in the foot who will require self-care. Based on our preliminary survey of two of the three blocks we anticipate there will be approximately 150 people in our cohort across 15 intervention areas. Each member of the cohort will have observations taken at four times at the time points defined in the schedule

Estimated precision

We propose to investigate several models for our analysis and select the best fitting alternative. However, we give an approximate estimate of the expected posterior variance based on the priors stated in the previous section. The average posterior variance is:

$$E\big(Var(\gamma|Y)\big)$$

where Y is the study data and the expectation is taken with respect to the prior distributions. For this analysis we assume no covariates and a linear function of time. Based on the priors stated above the expected posterior variance for the intervention effect in the prevalence model is 0.090. For the ulcer area model the value is 0.059. These values imply 95% credible interval widths of approximately +/- 0.59 and +/- 0.46, respectively. The smallest effects of the intervention for which these values would provide a 95% credible interval excluding zero would be odds ratios greater than 1.80 or less than 0.55 for prevalence, and greater than 1.58 and less than 0.63 for ulcer area.

Analysis plan of qualitative data

Quantitative data collected using educational test or checklists will be analysed using descriptive statistics.

For qualitative data we will use framework analysis guided by the MRC Framework.

We will code and analyse interview data for:

- Context - factors that influence or are affected by the intervention and its outcomes, and that prevent or enable change prompted by the intervention
- Intervention - how intervention is delivered and adapted.
- Mechanisms of impact: how do participants respond, what mediates this, and any unanticipated pathways and consequences.
- Impact of the intervention on the participants lived experience and that of the community.

Within our framework we will comparing data across sites and between data sources.

All interviews will be audio recorded and transcribed and translated into English. Observation notes will similarly be translated. Coding will be applied to all interview data whilst alert to new emerging themes and need for sub-themes.