Here is another response from Tom Pullum:

What I like to do is to try out any procedure that I suggest to someone else. In this case, you have done a lot of file preparation, and I cannot take the time to do the merging and recoding to get the data file you are working from.

The weight for the cluster should be the total weight of the children in the cluster, just as you said. You can get that total weight by defining wtd_n=1 and then collapse (sum) wtd_n [pweight=weight], by(whatever). That will give you the sum of the weighted cases If you want an unweighted sum then you could have unwtd_n=1 and then collapse (sum) wtd_n, by(whatever). If you want to combine weighted and unweighted sums on the same records, I think the easiest way is to do those two runs and then sort and merge on "whatever".

By the way, you should not use aweight. It is almost never appropriate.

My bigger question is this: why are you aggregating the data so much? If you know where the intervention occurred, just use that as a binary variable at the lowest level possible. The haz score, for example, is an outcome at the level of the individual child. I would do the analysis at that level, and not collapse at all. I strongly advise against aggregating when you do not have to. You may be able to use a multilevel model, i.e. a mixed effects model, with the children as level 1 and the village as level 2. Have you tried to do that?

In any case, this looks like an interesting project. Let me know if you have other questions.