

---

Subject: Re: Combining DHS and SPA data set

Posted by [Bridgette-DHS](#) on Mon, 16 Oct 2017 13:52:40 GMT

[View Forum Message](#) <> [Reply to Message](#)

---

Following is a response from DHS Senior Research Associate, Lindsay Mallick:

Since we cannot know the exact facility a respondent uses, we must link geographical groups (clusters or regions) of individuals to the most geographically proximate group of facilities. The size of the geographic areas or "buffer zones" depends on whether the SPA was conducted as a census or sample of all the facilities in the country and, if the SPA was a census, whether the facility is urban or rural. If it is a census, the buffer zones must also be large enough to capture the displacement of GPS coordinates for individuals, which is done to protect anonymity of the respondent. You would then summarize all the facilities in that buffer zone with a summary statistic of your choice (ie mean, median, etc...), and then merge those summary scores with the appropriate level (cluster or regional) in the individual-level file. More detail can be found in the methodology section of our Analytical Studies 51 or this report by Skiles et al., 2013.

---