Subject: Sampling Error In Light of DHS Analytical Report 46
Posted by shujaat.smc@gmail.com on Wed, 20 Jan 2021 13:44:34 GMT
View Forum Message <> Reply to Message

Dear Tom Pullum, Greetings & Good Morning!

In response my one of the posted question on the forum (House Vs Home Delivery) where I was confused about the insignificant association of the interventions for preventing maternal & neonatal death you shared an analytical report 46.

Where three possible reasons were mentioned by you,

1. The coverage of intervention was not complete.
2. Quality of healthcare provided at the facility may not be up to standard
3. False statement by respondent eg; they respond complete adherence to iron supplement but in reality, they were not.

Today I was reading a paper "How much DHS Data is Reliable" with a focus on neonatal mortality, copy attached.

In this paper on page, 17 were sampling error is discussed for the wide confidence interval and the author mentioned that (in my paraphrasing) the reason for wide CIs;

1. Number of neonatal deaths & the total number of birth recorded in the survey was lower. (Curtis 1995)

2. Relative SE can be very high if the country has low NMR

3. Relative SE can be very high if the sample size is relatively small.

4. Sampling Error increase with sub-group analysis.

Key Word for my question "Sample Size was Relatively Small & Subgroup Analysis"

I found that the number of women of reproductive age in Pakistan census 2017 (from which sampling frame was derived for fourth PDHS 2017-18), is quite high and the number of ever-married women who responded in PDHS 2017-18 was 12,364.

15-24 years: 19.3% (male 23,119,205/female 21,952,976)
25-54 years: 34.7% (male 41,589,381/female 39,442,046)
https://www.indexmundi.com/pakistan/demographics_profile.htm l

Quest; This number of ever-married women "12,364" who responded to PDHS 2017-18 is very low when we compare to the number I have
shared above for age 15-24 & 25-54.

I want your guidance that on the basis of such vast difference in the absolute number of population of women of reproductive age & Sampled women in PDHS 2017-18, this perfectly fits into the third statement "Sample Size was relatively Small". Hence the CIs in my analysis on the predictors of neonatal mortality become wide and world-famous predictors become insignificant in my model.

Therefore I got the fourth & Fifth reasons in addition to your three reasons mentioned in the Analytical Report 46. "The Sample size was relatively Small" and since I have limited the analysis to last born neonates (my sub-group).

Does my understanding is acceptable to the science of statistics & DHS methodological experts?

Looking Forward to your reply.

Best Regards

Dr. Hussain

## File Attachments

1) The_measurement_of_neonatal_mortality_How_reliable.pdf, downloaded 370 times

---

Subject: Re: Sampling Error In Light of DHS Analytical Report 46
Posted by Bridgette-DHS on Wed, 20 Jan 2021 20:03:08 GMT
View Forum Message <> Reply to Message

Following is a response from DHS Research & Data Analysis Director, Tom Pullum:

The relevant numbers for calculating the standard errors are the number of births (the most recent birth in the past five years) and the number of neonatal deaths, which are the denominator and numerator of the NNMR. The design effect (the reduction in effective sample size due to the sampling design) plays a role, in the direction of increasing the standard error. If you are using logit regression in Stata with svyset and svy, then your standard errors and confidence intervals are being calculated correctly.

The population numbers have nothing to do with the standard errors. The number of births in the sample is proportional to the number of women in the sample, but otherwise the number of women is not relevant to the calculation.

If you have a larger sample, you have more power, and are more likely to detect a difference that exists in the population. In general, if you do not get statistical significance, it MAY be because you don't have a large enough sample. It may be because of measurement error or reporting error, or because the relationship simply does not exist in the population. Those are pretty much

the only explanations.

Subject: Re: Sampling Error In Light of DHS Analytical Report 46
Posted by shujaat.smc@gmail.com on Sat, 23 Jan 2021 07:54:47 GMT
View Forum Message <> Reply to Message

Dear Tom Pullum Greetings!

Thanks a lot for your elaborated and concept clearing reply.

Best Regards

Dr. Hussain