

## POPULATION ESTIMATION BY CROWDSOURCING IN DHOK HASSU

### INTRODUCTION:

In public health, it is often necessary to have reasonably accurate estimates of populations of interest such as catchment population of a facility or in a target locality, children in an area etc. In developing countries such as Pakistan, such information is often not readily available. For e.g. the last completed census was conducted in 1998 and the latest one was conclude and its results are still not in public domain. In any case, there was a gap of nearly 19 years between the two events, while the country has seen both massive increase in population and internal and external migration which makes it extremely difficult to have meaningful estimates of populations of specific localities for public health interventions such as how many children to vaccinate.

In establishing our “urban laboratory” in Dhok Hassu, Rawalpindi, we gathered information from existing sources. We were told by public officials to anticipate a population of 65-70,000 in the 3 union council area comprising Dhok Hassu. On the other hand, UNICEF and WHO – that support vaccination activities - estimate that the population of households with at least one child under the age of 5 years of age was around 140,000. No information was available about the households with only older children, those with no children and households with single adults.

Our group has used crowdsourcing to estimate populations of localities and of healthcare providers previously. However, the concern has been about scientific rigour of the methods. In the estimations for Dhok Hassu, we divided the local population into lanes or spots. Of the total spots, around 57% were included in crowdsourcing and of these 102 (17% of those included in crowdsourcing and around 10% of the total) were verified by a household census.

### METHODOLOGY:

Using Google Earth, we identified 1040 “spots”. Each spot corresponds to an identifiable street of approximately 50-70 meters. In reality this reflects the smallest street that is visible on Google Earth and can accommodate an automobile. There are smaller “galian” (lanes) that only allow passage of pedestrians/ cycles that are not visible on Google Earth.

Out of the total spots, 597 spots were randomly identified and visited for “crowd sourced mapping”. We wanted to capture as many spots as possible with budget confines. In each spot, the team asked pedestrians or shopkeepers what they thought were their estimates of the minimum and the maximum number of 1) number of houses, 2) number of households and 3) total population of that spot. Between 5 and 7 informants were queried in each spot. Once the data were compiled, a median of all minimum values in a spot and a median of all maximum values for each spot were estimated. For each spot these medians were averaged to arrive at

### CONDITIONS REQUIRED FOR ACCURATE ESTIMATIONS FROM CROWDSOURCING

- There is no prior contact between interviewers and interviewees
- There is no prompting by interviewers
- The different interviewees must not confer with each other
- The medians of individual spots are estimated and then means of the medians are calculated
- Within spot variations can be very large but with sufficiently large number of spots, the overall estimate becomes increasingly accurate<sup>2</sup>

the estimated population of the spot. In doing so, we followed the methodology described previously.<sup>1,2,3</sup>

Within the 597 spots identified above, 100 spots were randomly assigned for a census. A separate team visited these spots and listed members from each of these households along with their ages.

Population estimates for crowd sourced mapping were compared with those from census for an estimation of accuracy of the crowdsourcing process. The overall process took around 7-8 days for the mapping and the costs were under Rs. 200,000.

## RESULTS:

Estimates of Crowd sourced Mapping and Census are accurate within 6%; with greater variation for individual spots. The total estimated population of Dhok Hassu 235,857 based on crowd sourced mapping. Within the spots where both census and estimation was conducted, the difference in the overall population was 1.8% higher from mapping than census based estimates. Within individual spots, the variation was much larger, as has been described.<sup>2</sup> Individual mohalla (neighbourhood) populations are as below:

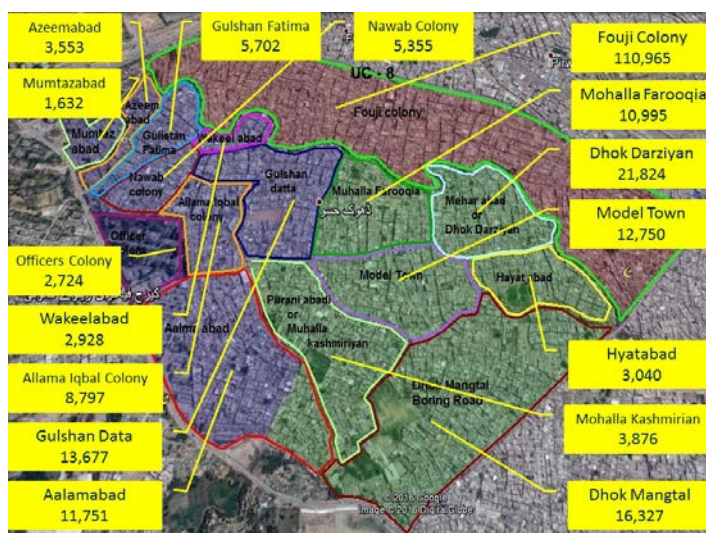
## INFERENCE

Crowdsourcing is a viable method to arrive at publicly known information. We applied this technique for estimating the overall population of a community. While the technique is based on certain assumptions about interviewees – such as local residents and that they do not confer with each other prior to answering our interviewers - these conditions are easily met. The estimates reached from the technique are more accurate than have been generally appreciated. We propose that the technique has wider applicable in public health and policy than previously appreciated and can serve as an interim strategy until more formal techniques such as the census become available.

TABLE 1: MOHALLA LEVEL POPULATION ESTIMATES

	Mohalla	Estimated population	Distribution
1.	Aalamabad	11,751	5%
2.	Allama Iqbal	8,797	4%
3.	Azeemabad	3,553	2%
4.	Dhok Darziyan	21,824	9%
5.	Dhok Mangtal	16,327	7%
6.	Fauji Colony	110,965	47%
7.	Gullistan-e-Fatima	5,702	2%
8.	Gulshan Data	13,677	6%
9.	Hayatabad	3,040	1%
10.	Model Town	12,750	5%
11.	Muhalla Farooqia	10,955	5%
12.	Muhalla Kashmirian	3,876	2%
13.	Mumtazabad	1,632	1%
14.	Nawab Colony	5,355	2%
15.	Officer Colony	2,724	1%
16.	Wakeelabad	2,928	1%
	<b>Total</b>	<b>235,857</b>	

FIGURE 1: MOHALLA LEVEL POPULATION ESTIMATES



<sup>1</sup> Galton F. Vox Populi. *Nature*. 1907; 75(1949):2.

<sup>2</sup> Surowiecki J. *The Wisdom of Crowds*. Double Day; 2004.

<sup>3</sup> Ranard BL, Ha YP, Meisel ZF, et al. Crowdsourcing-- harnessing the masses to advance health and medicine, a systematic review. *J.Gen. Intern. Med.* 1/2014 2014; 29(1): 187-203

For more information please contact



www.resdev.org  
Phone: +92 51 843 6877

Adnan Khan adnan@resdev.org  
Ayesha Khan ayesha@khans.org