Creating Household data: a conundrum of feasibility

Background:

Data collected by UNHCR at registration (PG henceforth) and during home visits (HV henceforth) provides case-based information. As we have come to realize, many cases live with other cases under the same roof – that is they live in the same house.

The HV questionnaire (both version 2 and the latest version 3 from which VAF data is collected) includes a question on "UNHCR file numbers for all members living in the house (in a different file number)". From this particular articulation comes the idea and opportunity to create 'household' data. This process has however highlighted numerous limitations and difficulties, as well as presenting an issue of definitions.

Concern 1: What is a Household?

WFP's definition of a household are "people who eat from the same pot"[1]. Families/cases in a 'household' live and pool together their resources to support one another. Whereas the definition of a 'House' is those living under the same roof. The difference may appear small but the implications are significant.

The HV form presents the opportunity to link cases that live together. This however creates 'house' data, not 'household' data per se. Different cases can live together but not "eat from the same pot". Living under the same roof reduces per person rent, utility savings etc. Using 'house' data (which is a summation exercise, where one aggregates case data) as 'household' data, requires the assumption that a pooling of resources strategy is occurring. This assumption is not a truism, and in instances where it is not, it presents a huge and inadmissible exclusion error (see pager "Case-based vs. House-based data: for a more efficient utilization and distribution of scarce resources" and annex 1), where highly vulnerable cases risk being cut from assistance. At present, giving the current HV questionnaire only a 'house' level view can be created, whereas a 'household' level view cannot.

Concern 2: Challenges of incomplete data

The second limitation to generate 'house' or indeed 'household' level data that has come to light is incomplete information. As of January 5, 2014 we have about 15,000 cases reviewed with the HV3 form. This is approximately 12% of the required home visits to reach census level data.

Of those 15,000 cases we know who they live with, but we do not have information on the majority of cases that they live with yet. Therefore we are unable to aggregate data completely. Only once we have comprehensive data will we be able to create a vulnerability picture of the entire house. To achieve census level data would require in the region of a further 100,000

---

[1] A standardized definition of 'household' across the VAF steering committee members and humanitarian agencies has not yet been agreed on and a methodology to collect this data has yet to be agreed.

home visits which would cost approximately $10 per review. As part of a cost benefit analysis we can calculate the potential cost vs. savings achieved through completing the home visits[2].

Concern 3: Where are you moving?

There are two issues relating to the static nature of data collection, each with its own implications. These are relating to family mobility and family composition.

Relating to mobility it is common that families move location. If the family has been visited twice, once before and once after a move, then we can simply use the most recent record for that case and house structure. However, there is a small number of instances where the date of the home visit has not been recorded. We have no way of knowing which is the most recent visit.

Relating to composition, within a family structure there is the possibility that the structure will change over time since births, deaths, marriages, cases moving, new cases joining a house, all account for additions or subtractions from house data. If house data is to be utilized for targeting and the composition of a house has changed we must re-estimate and re-interview the different cases in order to re-estimate their vulnerability (see annex 2).

For both instances there must be the acknowledgement of this data limitation but little else can be done to improve it.

Conclusions:

1) As of today, house level aggregated data is not possible until census data is achieved.
2) It is preferential to inform targeting based on household level data rather than house data. Case-level (as current) is the next best alternative to household level.
3) Household data, defined as "people eating from the same pot" who pool resources together is not possible with the data we possess. HH data can only be created if the data is collected in this specific manner. The survey must be HH-based and enumerators interview the head of HH rather than the head of case.
4) We must acknowledge limitations caused by data collection latency and also ensure that visit date is captured with each home visit.

Recommendations:

To alter the data collection form and to achieve census level data IF the cost benefit analysis is positive.

---

[2] A cost benefit analysis on a methodology for collecting comprehensive data across the refugee community will be conducted as part of the VAF project (during the first quarter of 2015).

Annex:

Annex 1 (Concern 1)

House formed of:	Expenditure of Case
Case A (1 person)	200
Case B (1 person)	100
Case C (1 person)	40
Total (per person):	113.33

Deprivation of Assistance: given a poverty line of 68, Cases A and B are non-poor, whilst Case C is poor. The house is non-poor (113.33 expenditure). Targeted assistance would exclude this non-poor house. However, they are not pooling resources together, as a consequence Case C is being deprived of much needed assistance.

Efficient Assistance: no need to provide assistance to 3 Cases in order to lift them out of poverty when you can provide assistance to only Case C and lift him/her out of poverty (as the rest of the house is already non-poor. Money is saved, and more resources are available to help more families in dire need.

Annex 2 (Concern 3)

Home Visit: January 2014		Home Visit: September 2014	
House formed of:		*House formed of:*	
Case A		Case A	
Case B		Case H	
Case C		Case J	

Issue: no Home Visit carried out for Cases B and C post September. We are aware these cases no longer live with Case A (old data is therefore incorrect), but we also don't know with who they live today. We have no way of creating House data until we re-interview Cases B and C (we can still use Case data, assuming Case-composition has not changed).