
Obtaining representative data on IDPs: challenges and recommendations

Contributed by: Natalia Baal and Laura Ronkainen¹

With the global estimates of persons displaced by conflict and violence reaching record numbers year by year – with those internally displaced (IDPs) making up the largest portion² – there is an increasing call for more accurate data on these persons' situations.³ The rising figures have been accompanied with calls for more attention to support states affected by internal displacement in addressing needs and finding solutions to displacement through humanitarian, development and peace building efforts. Increasing attention on displacement in policy making and programming has also led to a stronger focus on obtaining and using sound evidence on displacement situations as the starting point for advocacy and response. Representative and valid data, generated through statistical methods rooted in scientific good practice, is thus more and more called upon to provide decision-makers with objective guidance on where and how to spend resources.

The availability of and access to information from various sources and with more timely and regular updates have made the use of evidence for decision-making standard and good practice. The use of advanced research and analysis methods in displacement contexts is becoming increasingly part of day-to-day operational work as the practice of information management becomes professionalized among the actors delivering response. At the same time, use of the latest technology and finding new ways of collecting data such as social media analysis, aerial imaging and use of call detail records, have introduced different ways of extracting operationally relevant information in humanitarian contexts.⁴ Growing attention on the topic of displacement together with the increased availability of data has transformed what was a few years ago a repetitive call for 'more data' into a more advanced call for 'better data' that can provide statistical and comprehensive 'hard evidence' on the displacement situation. However, due to various operational, political and conceptual constraints, obtaining data that can be considered representative of the IDP population as a whole remains a quandary in most IDP contexts.

¹. Both working with the Joint IDP Profiling Service (JIPS). The views and opinions expressed are those of the authors and do not necessarily represent the views of the Office of the United Nations High Commissioner for Refugees.

². According to IDMC estimates, at the end of 2016 the number of IDPs displaced by conflict and violence totaled 40.3 million, after an estimated 40.8 million and 38 million at the end of years 2015 and 2014 respectively.

³. Including outcomes of the World Humanitarian Summit, the 19 September High Level Summit of Refugees and Migrants and a report by the World Bank Stock-taking on Forced Displacement Data, Oct 2016.

⁴. For example using call detail records for tracking trends of population movements over time such as in Bangladesh during and after cyclone Mahasen in project developed by the United Nations University Institute for Environment and Human Security, the International Centre for Climate Change and Development, Flowminder, Grameenphone and Telenor Group; or using satellite imagery in Mogadishu by UNOSAT and partners to map out areas and follow changes in urban morphology to understand the overall trends of intra-urban displacement and population increase in areas with high security risks.

This article will briefly discuss the technical prerequisites for obtaining representative data in IDP contexts and highlight some of the challenges that often stand in the way of achieving this. Some food for thought and recommendations will follow to suggest ways to reconsider the types of data and processes for obtaining them in IDP contexts. These suggestions are based on the experiences of the Joint IDP Profiling Service (JIPS)⁵ when providing support to partners working on data processes in various contexts affected by internal displacement.

What is needed for IDP data to be representative?

Simply put, representative IDP data means that when analyzed, the observed characteristics of the population emerging from the data reflect the true characteristics in the total IDP population that is being researched. This kind of analysis can be obtained either through data covering the total IDP population, or through a representative sample.

In most IDP contexts, no registration systems or other list of the total IDP population exist to enable analysis on their situation. Where they do exist, they are often incomplete due to the high cost of maintaining them up to date, as well as various other reasons that will be discussed later in this article. These gaps, often corresponding to groups likely to be particularly vulnerable, create biases that should be of concern to the humanitarian community. In the absence of such lists in most countries, data on the total IDP population is often not available at all, or consist of a patchwork of estimates from various sources collected for different reasons (most commonly a patchwork of uncoordinated programmatic assessments implemented by different actors, in different but overlapping places and time periods). Some examples of including displacement analysis in the national census or large scale demographic surveys to obtain an overview of the displacement situation exist from some countries, but in general the opportunity for collecting or analyzing information on internal displacement in these processes is rarely tapped into.⁶

Hence, most of the time data with complete coverage of the total IDP population is not available. Instead, a sample-based approach is often used to produce data with the objective of generalizing the findings to the total IDP population. Choosing a sample should allow for each individual in the IDP population, be it households or individuals, an equal opportunity to be selected for study. Sample-based methods are often the most cost-effective and feasible for obtaining needed data on IDPs, and – done well – they can produce equally useful information for operational decision-making as data collection with full coverage; a point that is far too often under-valued. Choosing a sample that can produce representative data on the IDP population requires careful planning based on information, at a minimum, on the locations and estimates of the size of the overall IDP population, which can be used as a sampling frame⁷. However, various issues complicate not only the establishment of a comprehensive sampling frame on IDPs, but also collecting data on the situations of these populations in general thus further complicating the ability to produce representative data.

5. JIPS is an inter-agency service providing profiling support to governments, humanitarian and development actors responding to displacement situations. It is overseen by an Executive Committee comprised on UNHCR, OCHA, UNDP, DRC, NRC-IDMC and the Special Rapporteur on the Human Rights of IDPs. More information about JIPS and the profiling exercises mentioned in this article can be found on the JIPS website: www.jips.org.

6. Research carried out by Vebjørn Aalandslid, Frode Berglund and Kari-Anne Lund from Statistics Norway seconded to UNHCR and JIPS through NRC/Norcap. See also Aalandslid, Lund and Berglund (2015): *The population census as a tool to capture statistical information on forcibly displaced persons*, UNHCR Statistical Yearbook 2014, UNHCR.

7. A sampling frame is a list of the people in the population studied that will be used for selecting the sample. With an incomplete sampling frame, analysis based on the drawn sample risks not representing the total population.

Challenges for collecting data on IDP populations

The challenges of designing and carrying out a data collection process that can produce representative data on IDPs stem from different sources, here grouped as operational, conceptual and political considerations. Some of these challenges are common to any data collection effort in humanitarian contexts or contexts with high human mobility, whereas others are more specific to situations of conflict- or violence-induced internal displacement.⁸

Quick and dirty? Operational challenges effecting IDP data

Collecting data on an IDP population clearly faces the same issues as any data collection effort in humanitarian contexts. In addition to the often-limited resources available for putting together a comprehensive map of where the IDP population is and what kind of situations they are facing, both crucial steps to developing an appropriate sampling design, the challenge is increased by restrained or complete lack of access to certain areas. Examples of such constraints include IDPs living in remote locations with poor infrastructure, such as in many parts of Colombia, where these obstacles make keeping a country-wide database of IDPs and other victims of conflict up-to-date challenging; volatile security situations such as in Mogadishu, Somalia, where some parts of the urban area had to be excluded from a household survey collected through a profiling exercise due to insecurity; and other unforeseen factors, such as the Ebola outbreak, which prevented the teams collecting the data to have access to some parts of the country for a national displacement profiling exercise in Côte d'Ivoire. Particularly in urban settings and other contexts where IDPs mix with other population groups, identifying them, whether for counting or a further analysis of their situation, is often challenging, if not impossible, even if physical access to the areas is not limited. Human populations are also rarely static, and fluid population movements, either forced or voluntary, often get ahead of the data collection teams, resulting in potential double counting or exclusion, as well as data that may be out-of-date before data collection even has finished.

Unrealistic timeframes between the decision to obtain data on IDPs and deadlines for decision-making based on the results often limit what can be achieved in terms of coverage and precision. Establishing protocols and systems for data sharing with partners in a timely manner, which would help avoid duplicated efforts, is often not an operational priority. In addition, existing data are often difficult to use as part of analysis due to ill-documented or incompatible methodology descriptions, definitions of locations and boundaries or even definitions of the target population as discussed in the next section.

Limited technical capacity of actors collecting data on IDPs both in terms of questionnaire design and understanding of the displacement situation can lead to the use of tools and methodologies that are inadequate for capturing the full picture of the displacement situation. An example all too often observed in the field is questionnaires where the self-identification of IDPs culminates in the binary question of: "Are you an IDP?" This can lead to 'false positives' if data collection is linked to assistance provision, or to IDPs choosing to hide their displacement-related identity if this is linked to fears of stigma or even persecution. To mitigate problems such as this, collecting reliable, comprehensive and coherent data on displaced populations requires various and specific capacities in the field of demographics, internal displacement and protection, as well as relevant contextual knowledge.

8. The authors recognize that many issues highlighted in the article also relate to displacement caused by disasters, as recently noted in discussions at the Platform on Disaster Displacement's Advisory Committee's workshop. However, some different operational, conceptual and political dynamics would also need to be included should disaster displacement be adequately addressed.

However composing such a team requires dedicated resources, communication between experts from different fields, and often reaching out beyond the capacities available in any one agency or department.

Beyond numbers, the analysis on IDPs' socio-economic situation, vulnerabilities and capacities often only find their meaning through a comparative analysis with other population groups living in similar situations. Linking IDP data collection to broader data systems by ensuring compatibility of used methodologies and tools from the start can be extremely helpful for the greater usability and longevity of the data. Working with national statistics offices is often a good way of ensuring this link, while also tapping into existing national capacity. However, where national capacity on IDP or population data management generally needs strengthening, or collaboration between the national and international actors requires specific administrative arrangements, firm deadlines for humanitarian decision-making, such as for funding cycles, often take precedent over longer-term investments in collaboration and capacity building. This tendency limits the potential for the data to have a greater impact through a nationally-owned process.

Who counts? Conceptual challenges of IDP figures

Operational considerations are not the only obstacles for obtaining sufficiently representative data on IDPs to form a basis for analysis; in fact, the issues go much deeper. The Guiding Principles on Internal Displacement's definition of internally displaced persons is an inclusive one, but in practice actors collecting data on the displacement situation on the ground often draw the lines narrower in order to focus limited resources on those IDPs that response is targeted for. In Mogadishu, Somalia, for example, the working group leading an urban displacement profiling had lengthy discussions on how to draw the limits of the scope of the exercise in a context where, according to the Guiding Principles' definition, more or less everyone in the city would qualify as an IDP. To keep the process manageable and to produce data that would be useful for operational purposes, it was agreed that the profiling – and the resulting updated IDP figures – would only focus on the numerous settlements where IDPs were known to reside in sub-standard living conditions.

Deciding on whom to count and who counts as an IDP present other challenges that need to be considered when planning for a data collection exercise aiming for representative data. Many data systems producing estimates or figures on IDPs are based on the logic of counting population movements. While these systems are providing extremely useful information on the trends of population movements over time, using these flow figures as a source for estimates on the total number of IDPs brings various problems. Internal displacement does not have a specific point that makes the move count as 'an IDP movement' the way that crossing an internationally recognized border does. Displacements for short distances, even just half a kilometer, can result in similar tragedies as displacement to the other side of the country: loss of livelihood, land, documentation or even a family member. However, these short-distance moves that often take place within one urban area and for small numbers of IDPs are easily omitted by monitoring systems. Furthermore, the detrimental impacts of displacement to IDP individuals, households and communities may be just as grave even if the physical absence from the place of habitual residence only lasts a short period of time. These incidents thus also merit acknowledgement in displacement figures and analysis.

Although difficult, data collection methods with broad coverage and significant resource investment can produce relatively good estimates of those populations that were once forced to flee their homes. Once obtained, however, keeping these figures up-to-date for life events such as births and deaths poses a further challenge, such as in Colombia, where despite considerable efforts the registration system has a persistent backlog. Conceptual challenges are troubling these processes as well: if births

to IDP families are recorded in displacement figures, for how many generations should this continue?⁹

An even harder nut to crack for data collection is when to consider a person once categorized as an IDP as no longer an IDP. IDPs establish new lives and find solutions to displacement-related challenges. While the Interagency Standing Committee's Framework on Durable Solutions for Internally Displaced Persons does provide a definition on a durable solution to internal displacement and work is underway to operationalize this to measure overall progress towards solutions,¹⁰ the complex nature of the topic makes establishing a foolproof, single tipping point from displacement to a solution equally applicable across all contexts arbitrary. Although this arbitrariness can be lessened with an appropriate evidence-base for informed decision-making, the decisive element cannot be completely avoided.¹¹ And finally, IDPs' voices through the collection of subjective data also need to be heard in defining when and how they have found a solution – an ideal that is rarely realized in data collection that attempts to simply get the figure right.

Interests and identity: Political challenges surrounding IDP data

Conceptual challenges can foreseeably be overcome with a concerted effort to clarify and systematize key questions, however internal displacement caused by conflict or generalized violence and related figures and data are all too often highly politicized making this effort nigh impossible to translate into practice on the ground.

International and national actors alike have varying interests that underline a certain discourse around IDP figures: maintaining the status quo for legitimacy and continuity of response systems and resources in place, or changing it by inflating or deflating the figures. Moreover, these factors can play out differently at national and sub-national levels, becoming intertwined in political discourse from community, through municipality, up to ministerial levels. Consequently, discussions on IDP figures are often caught in an impasse of one reading of the figures against another, and however rigorous the statistical analysis that is conducted, the results will categorically be rejected by some actors unless consensus on a joint and shared approach to the data process is agreed upon from the start.

But displacement is not only political from the side of national institutions and international organizations delivering response, it also has a human element. In the end, as registration systems with adequate 'IDP identification mechanisms' rarely exist, data collection exercises rely on people self-identifying themselves as IDPs or non-IDPs. For practical and socio-political reasons, the displaced and non-displaced populations often opt themselves in or out of the count for various reasons. How often this happens is almost impossible to tell. If identifying oneself as an IDP is clearly linked to assistance provision that non-IDPs will not benefit from and there are no substantial risks associated with being identified as an IDP, it can obviously be expected for a few non-displaced persons to opt in to the list. On the other hand, IDPs may also not wish to be identified as such, for example if tensions between the host community and IDPs are tangible due to limited resources, such as in the Kurdistan region of Iraq; or if an IDP has fled to hide himself from a gang threatening his life, such as in Honduras. Hidden – or hiding – IDPs are a methodological challenge, but collecting data on them is

9. See also IDMC (2014): *Born in displacement: Challenges in assisting and protecting descendants of internally displaced people*.

10. Under the leadership of the Mandate of the Special Rapporteur on the Human Rights of IDPs and in collaboration with a Technical Steering Committee consisting of a broad group humanitarian, development and peace-building actors, JIPS is implementing a project to operationalize the IASC Framework into measurable indicators, tools and methodologies to measure progress towards durable solutions to internal displacement. For more information, please visit: <http://www.jips.org/en/profiling/durable-solutions/measuring-progress-towards-solutions>.

11. Different efforts are currently underway in both refugee and IDP settings to develop integration indexes and measurement tools to help inform programmatic and operational decision-making; each one is facing this specific challenge and managing the result in different ways.

also often a decision with important political and protection implications. In some cases, there are ethical issues involved regarding the collection and dissemination of statistical data if it risks aggravating existing tensions and drawing unwanted attention to groups that may wish to remain hidden.¹²



Conclusions and food for thought

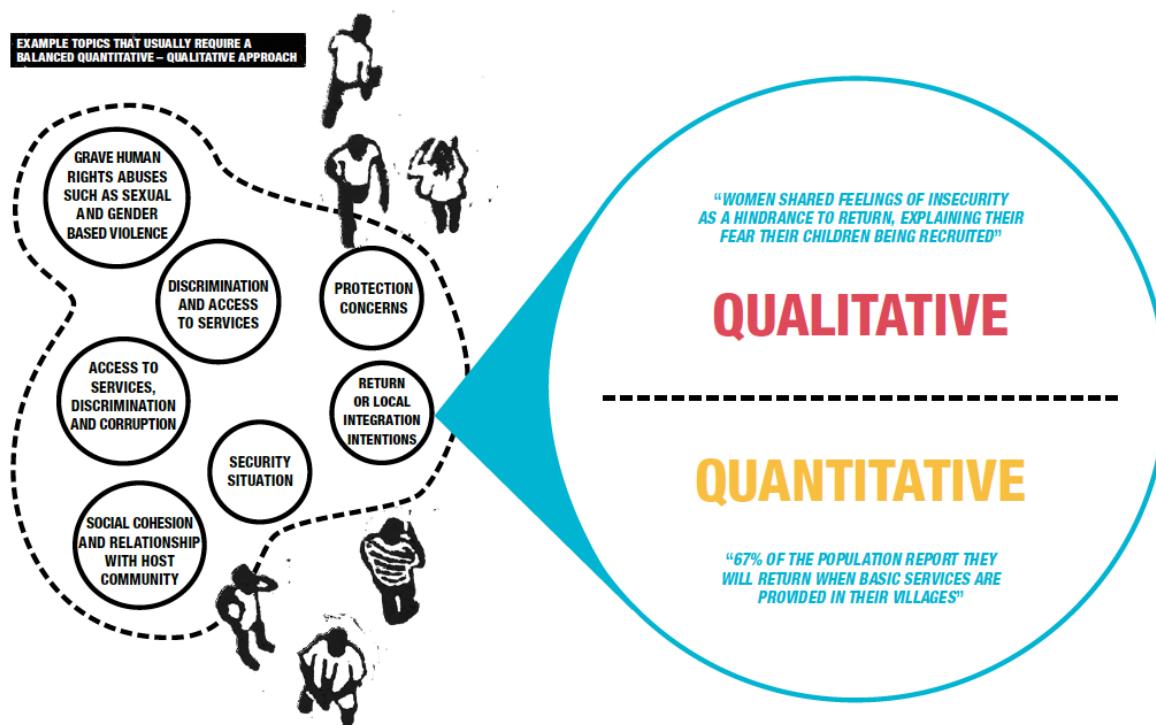
This article has introduced a plethora of operational, conceptual and political challenges that make collecting representative data on IDP populations difficult. Although many of the impacts of these challenges on the resulting data can be minimized through appropriate actions including adequate planning, capacity building and negotiation, and others can be mitigated through embracing new technologies and utilizing improved methodologies, an important few are less easily addressed by the very nature of the phenomenon they are trying to capture.

Rather than detailing a to do list of mitigating strategies that exist in many guidance documents and trainings to address the first two groups, this article will close by sharing a few broader reflections and recommendations that have proven (and are likely to prove) useful for overcoming some of the more challenging issues presented earlier. These relate to collaboration and consensus building, strengthening international standards, the use of qualitative data, and, most importantly, repeatedly answering the basic question of ‘how will this data be used?’.

12. An ongoing process to develop the discipline of protection information management is here particularly relevant as it sets out key principles and helps to operationalize them, including protection risks, to ensure protection is mainstreamed into data collection, processing, analysis and dissemination.

Collaboration on data processes can help reduce the amount of resources spent on data by different actors. Agreeing on common terminology, definitions, indicators and compatible tools is also a way of ensuring that the data resulting from different processes is useful to a broader group of actors. This can also help in unlocking some of the challenges mentioned around political sensitivities of the data: while quality data and evidence are needed as the foundation for sound decision-making in displacement contexts, often the key step in effective joint responses to displacement is agreement on some of the key elements of this evidence. We would argue that agreeing on the lowest common denominator on data and using this for joint responses can be more effective than datasets obtained even through the most rigid of statistical methodologies if some of the challenging questions highlighted in the above sections have not been answered by the actors together first.¹³

Answering these questions in country A or country B would be greatly helped (although not completely resolved for reasons described above) if there were clearer guidance at the global level. In this regard the newly established UN Statistical Commission's Expert Group on Refugee and IDP Statistics (EGRIS) holds much potential.¹⁴ Although during 2017 work on IDP statistics will only take the form of a technical report to be submitted to the UNSC in 2018, this will likely pave the way for a formally endorsed handbook and result in clearer, more systematic guidance and recommendations for transforming the IDP definition and some of its cross-cutting conceptual questions into data.



13. A key feature of JIPS' identity and work as an interagency service is its focus on facilitating consensus around data through collaborative profiling exercises. The purpose of this approach is to increase the impact of data through a transparent and participatory process that results in widely endorsed findings and recommendations jointly developed by the actors responding to a displacement situation.

14. The UN Statistical Commission's Expert Group on Refugee and IDP Statistics is co-chaired by UNHCR, EUROSTAT and Norway. Its work on IDPs is led by JIPS, supported by IDMC and Norway, and benefits from membership of several displacement affected countries (including Azerbaijan, Bosnia & Herzegovina, Colombia, Kosovo, Kurdistan Region of Iraq, Somalia, Philippines, Uganda and Ukraine) and international experts from UNHCR, the World Bank and the office of the Special Rapporteur on the Human Rights of IDPs.

As a word of caution, the emphasis on representative data and statistics may distract from the more important questions of what data is actually needed for response, and what the most appropriate method of obtaining that data is. Even if representative quantitative data is often perceived of and praised as ‘hard evidence’, it can rarely produce all the answers, and needs to be complemented by qualitative approaches. Qualitative data on sensitive issues with a more nuanced insight into people’s perceptions of their own situation and their fears and hopes for the future, also collected and analyzed with the necessary rigor and skill, can produce equally important operational evidence. Focus on quantitative methods only often over-simplifies the issues and can result in reinforcing assumptions through predetermined answer options and question patterns, rather than allowing for the data and the people behind it to speak for themselves.

Finally, we would like to put on the table the question that should always be asked first in every data collection process: “How will the data be used?” Although seemingly obvious, in reality this is often inadequately answered and usually overshadowed by the discussion on how to obtain the most reliable data and the rush to roll up our sleeves and start collecting it. A balance needs to be found first and foremost between the resources used for data generation and the resources available for response; indeed, ethics dictate that we should not be spending all our funds collecting data that we cannot act upon and, at the same time that our actions should be appropriately informed. A balance is also critical when considering the planned use of the data and the types of data needed, as ‘interesting to know’ and ‘useful to know’ do not always fully overlap from an operational perspective. In addition, more data, however representative, does not help resolve some of the other common bottlenecks to understanding a displacement situation: the lack of profound and contextualized analysis of already-existing data, and the reluctance to share this data with partners, thereby limiting its use and impact.

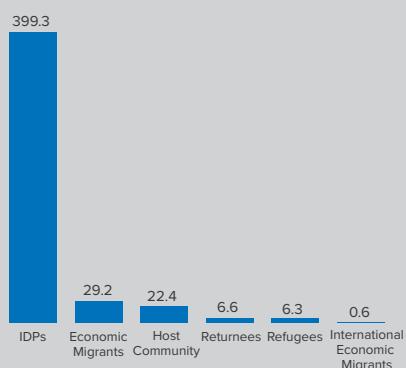
In summary, while an evidence-base is paramount to avoid misguided assumptions and inform effective responses, collecting data on IDPs faces significant operational, political and conceptual challenges that often make obtaining representative analysis difficult or even impossible. Out of these challenges, surmounting some should be prioritized to improve any data collection process – namely those linked to collaboration, choosing data collection methods that are fit-for-purpose and agreeing on common approaches to conceptually challenging aspects of IDP data at the global level when possible and in each displacement context as necessary. To overcome or mitigate others, significant investment of resources and time are required, and the resulting data may still be associated with caveats that will be of concern to the humanitarian community. Therefore, before launching an effort aiming to collect representative IDP data, it is recommended to stop first to consider whether a good enough, reliable and agreed upon compromise may well suffice for an equally effective response.¹⁵

15. For further reading on arguments presented in this article and others, please refer to Chemaly, Baal and Jacobsen (2016): *Forced Displacement: Go Figure! Shaking the Box of Profiling IDP Situations*, JIPS and Feinstein International Center. http://www.jips.org/system/cms/attachments/1174/original_2016-08_Forced_Displacement_WEB.pdf

Case study: Displacement profiling in Mogadishu, Somalia

Acquiring a complete and accurate picture of the IDPs and displacement-affected populations in Mogadishu is challenging due to the volatile security and fluctuating displacement situation. A collaborative profiling exercise was rolled out in 2015 to generate an empirical evidence-base to inform durable solutions strategy development and programmatic planning for addressing the displacement situation in the Somali capital. The profiling was a collaborative process coordinated by the IDP Profiling Working Group of the Protection Cluster, which consisted of local authorities, local and international NGOs, and UN partners, and was co-chaired by the Somali Federal Government's Disaster Management Agency. In particular, the profiling aimed to provide a disaggregated figure of the IDP population living in settlements in the city and provide analysis on their current situation, including needs and coping mechanisms, and preferences for durable solutions

**Distribution of the overall enumerated population by target group
(in thousands)**



The Profiling Working Group had to overcome several challenges to successfully attain the profiling objectives. At the beginning of the process, no up-to-date, agreed-upon figures of the IDP population living in the urban area existed. In addition to the dire security situation and on-going large-scale population movements, including intra-urban movements caused by evictions, obtaining an estimate on both the numbers of persons living in the settlements and their needs was also complicated due to the suspected purposeful inflation of the figures and exaggeration of the humanitarian needs of the populations in order to gain assistance by the settlement residents, members of the host

community, and so-called settlement gatekeepers. In general it was widely recognized that not all IDPs live in the settlements, and not everyone living in the settlements are IDPs, but with the settlement populations constituting a significant proportion of the total urban population living in areas with distinct humanitarian and development needs, the profiling thus focused on understanding the situation in these areas.

The profiling aimed to provide representative data of the different population groups that were identified in the settlements as a result of a comprehensive enumeration exercise, followed by a household survey. Due to limited access, the survey could not cover the two most insecure districts of the city. On-going movements challenged the exercise both in terms of sampling approaches, and in regard to the longevity of the data. The survey was also not able to collect representative data on all the different population groups living in the settlements, but focused on the groups with the highest numbers.

Despite these limitations, the profiling resulted in a comprehensive and broadly agreed-upon dataset and quantitative analysis of the displacement situation in the settlement areas, supported by qualitative data on specific issues. This analysis formed an evidence-

base for concrete recommendations on how to support sustainable solutions to all population groups living in the settlements through humanitarian and development interventions, including incorporating displacement into development plans and formulating a Government-led strategy on IDP solutions.

To read more about the profiling exercise and findings, download the full report at: <http://www.jips.org/en/field-support/country-operations/somalia/somalia> or get a snapshot through the Profile at a Glance (http://www.jips.org/system/cms/attachments/1262/original_JIPS-Somalia-Mogadishu-profile-spread.pdf). The quantitative data can be further explored here: <http://www.dart.jips.org/visualise?dataset=mogadishu>

