7

Counting the population in need of international protection globally

Brad K. Blitz, Alessio D'Angelo and Eleonore Kofman

Introduction

The significance of counting populations in need of international protection has become increasingly acknowledged as central to the effective design and delivery of humanitarian policies. Recent initiatives have emphasised the need for better data (IOM and McKinsey, 2018), as evidenced in the 2016 Global Compact for Migrants and Refugees and the 2030 Agenda for Sustainable Development. Of the 17 goals and 169 targets included in the Agenda, one specific migration target seeks to 'facilitate orderly, safe, regular migration' while others concern legal identity, birth registration, human trafficking and disaggregating data. Thus the International Forum on Migration Statistics held in Paris in January 2018 noted the salience of public debates around data and the need to look at data in the broader context within which they are created.

According to the UN High Commissioner for Refugees (UNHCR), by the end of 2016 there was a total of 22.5 million refugees,[1] with 17.2 million under UNHCR's mandate and 5.3 million refugees registered by the United Nations Relief and Works Agency for Palestine Refugees in the Near East (UNRWA) (UNHCR, 2017). The total number of refugees is the highest on record, although the annual rate of growth has slowed since 2012. In 2017, approximately 2 million people lodged applications for asylum while 2.8 million asylum seekers were awaiting determination of their refugee status (Table 7.1)**[[ok?]]**.

Table 7.1: Number of first instance applications lodged in 2016

Host state	Number of applications
Germany	720,000
United States	262,000
Italy	123,000

UNHCR estimates that, at the end of 2016, those under 18 years of age constituted roughly 51% of the global refugee population. The proportion of adult females remained relatively stable, at between 47% and 49%, over the same period. An estimated 60% of refugees were located in urban areas at the end of 2016.

However, one major problem complicating the task of effective humanitarian protection is the absence of accurate data on the populations most affected. For only some of the populations are disaggregated data available by age and sex. Furthermore, official statistics and data from major international agencies often struggle to cope with the messiness and fluidity of categories of people on the move as they transit, reside and possibly return to countries of previous residence or origin. Data are compiled by a diversity of organisations ranging from national states, regional bodies and international organisations, many of which operate with different definitions or collect data with different levels of detail and disaggregation.

This chapter reviews statistics and the coverage of those in need of international protection as set out in the UNHCR's guidelines and the different institutions involved in data collection. It identifies gaps in datasets used by UN and multilateral agencies tasked with the protection of refugees, internally displaced people (IDPs) and other people of concern and suggests that these need to be broadened to include other categories of vulnerable individuals and groups and that further disaggregation is needed.

Sources of data

Data sources on refugees and people in need of international protection are compiled by states as well as a range of international organisations including the United Nations Department of Economic and Social Affairs (UNDESA), the Organisation for Economic Co-operation and Development (OECD), the International Labour Organization (ILO), the World Bank, the United Nations High Commissioner for Refugees, the Internal Displacement Monitoring Centre (IDMC), the United Nations Office on Drugs and Crime (UNODC) and the International Organization for Migration (IOM).

Of particular relevance are the datasets produced by the UNHCR. While other agencies also deal with migrants, the UNHCR is the most authoritative agency mandated to protect refugees, asylum seekers, IDPs, returned refugees, as well as stateless people. For some populations – mostly in developing countries – UNHCR has additional demographic and socioeconomic information, including

date and place of birth, language, occupation, civil status, religion and education level. In locations where governments are exclusively responsible for data collection, comprehensive disaggregated data on refugees, IDPs and others of concern are often lacking or unavailable.

In addition to the UNHCR, we note the growing importance of the IOM as a source of data and in particular its establishment of the Global Migration Data Analysis Centre (GMDAC), jointly with the Economist Intelligence Unit (EIU) and development of the first global Migration Data Portal, which includes information on immigration and emigration; migrant flows; vulnerability; integration and well-being; forced migration; development; migration policy and public opinion. Other UN agencies such as UNICEF**[[please give in full]]** gather significant information which is published annually in the State of the World's Children Report, including data on child migration and displacement as well as statistics which bear on post-conflict settings and which are relevant to international protection, for example information on birth registration. The UN Office on Drugs and Crime produces the Global Report on Trafficking in Persons which includes data on victims of trafficking as well as projections for undetected victims.

Most of these global datasets are dependent on figures and estimates provided by national statistical offices, sometimes collected by specific agencies including police and law enforcement. These departments provide baseline and crucial demographic information, much of which is drawn from life history events and regular censuses. In addition, there are cyclical and sector-specific surveys which gather information on behalf of national governments and international agencies. For example, the Multiple Indicator Cluster Surveys (MICS), household surveys developed by UNICEF, seek to provide internationally comparable and statistically rigorous data on the situation of women and children. Since it began in 1995, more than 300 surveys have been conducted in over 100 countries, measuring the progress of the Millennium Development Goals and now Sustainable Development Goals (SDG)**[[ok?]]**.

The MICS are frequently compared to the Demographic and Health Surveys (DHS) initiated by the US Agency for International Development (USAID), which collects information on fertility, reproductive, maternal and child health, immunisation and survival, HIV/AIDS; maternal mortality, child mortality, malaria, nutrition among women and children (see also Chapter 6, by Roy Carr-Hill). One important difference is in terms of geographical coverage. Related to the DHS is the Key Indicators Survey (KIS), which supports the

evaluation of US government programmes in population and health in developing countries. The value-added of this dataset is that it can be used to produce household-level data for small areas including regional and sub-regional units as well as producing nationally representative surveys.

There has also been an increase in the amount of data collected by regional organisations, especially the OECD and European Union. The OECD is now among the most authoritative sources on the mobility of labour and presence of foreign-born nationals, though their datasets do not include refugees and asylum seekers as defined categories and foreign-born populations do not reflect immigration status or policy categories (for example students, highly skilled migrants, or refugees).

Eurostat, the statistical arm of the EU, coordinates, collates and systematises national data from member states, including the presence of asylum seekers in European host countries. It has provided data since 2008 under the provisions of Article 4 of Regulation (EC) 862/2007. Data are provided to Eurostat monthly (for asylum application statistics), quarterly (for first instance decisions) or annually frequency (for final decisions based on appeal or review, resettlement and unaccompanied minors). These statistics are based on administrative sources and are supplied by national statistical authorities, home office ministries/ministries of the interior or related immigration agencies in the EU member states.

Two different categories of persons are available in the Eurostat databank. The first includes asylum seekers who have lodged an application under consideration by a relevant authority. The second is composed of recognised refugees, or those granted another kind of international protection (for example subsidiary protection or authorisation to stay for humanitarian reasons), or those whose claims were rejected. The rates of recognition under each of the categories vary enormously for the same nationality in different EU member states. Since the entry into force of Regulation (EC) 862/2007, statistics on asylum decisions have become available for different stages of the asylum procedure.

Additionally Frontex, or the European Border and Coast Guard Agency, was established in 2004 for the management of operational cooperation at EU external borders.

Nongovernmental organisations (NGOs) and research bodies also produce independent datasets and occasional surveys which, while they may offer more specific coverage, still provide useful information. For example, the most widely publicised dataset produced by Amnesty

International and the polling company GlobeScan took the form of a global survey which used rankings to explore the degree to which respondents in the selected states were 'willing' to let refugees live in their countries, towns, neighbourhoods and homes, among other welcome indicators (see Blitz, 2017).

Limitations

Data collection in the context of conflicts, violence and disasters is inherently challenging (Abel and Sander, 2014; Bakewell, 1999). Methodological, operational and political complexities arise due to the nature of emergencies, as well as the characteristics, vulnerability and 'visibility' of the populations being measured (Chatelard, 2010).

The shortage of quantitative data on environmentally induced migration represents a key gap. The multi-causal nature of such movements also presents particular challenges and increases the risk of over-counting people on the move, or failing to capture them in the dataset altogether. Evidence on long-term trends associated with disasters, such as duration of displacement and subsequent movements, is also scarce (IOM and McKinsey, 2018**[[changed to match reference]]**).

Two key limitations are the coverage of populations of concern and the definitions used in sampling selected groups, which raises the question of possible biases in the data-gathering process. While people may enjoy different statuses under law which bring with them varying entitlements, these labels have become politicised (Zetter, 2007) and there are often competing interests at stake between donors and host providers (Harrell-Bond et al, 1992; O'Donnell, 2017**[[please provide reference]]**). States are under pressure in many instances to reduce numbers of refugees and asylum seekers on their territory, by removing them or closing down camps. In other situations, there may equally be a strong incentive for inflating numbers of people since that may inform the amount of aid directed to the host state (Edwards, 2013).

Most importantly, the terms used to describe those in need of protection bear on their entitlements, including levels of financial assistance received and opportunities for integration, including the right to family reunification. In practice, recognised *and* non-recognised refugees, asylum seekers and economic migrants may be at risk of abuse, including trafficking, and may require humanitarian protection. The critical issue is under what category the protecting authority decides to place them.

Further, although UN agencies and partner organisations such as the World Bank are among the most authoritative sources, their data, however similar, are not always identical. In practice, coverage may differ based on the organisation's mission and rationale for collecting data. One consequence of such gaps is the challenge of looking for causal relationships.

The gaps between such datasets are especially important in the context of ill-defined groups of displaced people where there is much potential overlap. We note that the UNHCR's definition of people of concern departs from some others which do not include those who are returned (UNHCR, 2016**[[please provide reference]]**). It is especially difficult to gather accurate data on returnees. Similarly, the quality of UNHCR's data varies from one category of protected person to another, thus potentially raising questions for the management and delivery of protection-related services. Most evident is their reporting on stateless people.

While the UNHCR has data on registered refugees for most countries, their data on stateless people and IDPs are incomplete. Even though the UNHCR has required its country offices to include stateless people among those to be monitored, it did so with little guidance regarding the definitions of certain categories of stateless people. As a result, it has published reports suggesting that in some large countries such as Brazil there are just four stateless people.

A report published by the UNHCR in 2010 advocated including de facto stateless people – those who may have no effective tie to their country of nationality - under existing categories (UNHCR, 2010). Elsewhere, the UNHCR has tended to aggregate data, bracketing larger numbers of 'people of concern' who did not fall neatly within legally proscribed definitions based on UN conventions and interpretations of international law. This problem is not exclusive to the UNHCR; elsewhere, others have noted the ambiguities of portmanteau expressions, such as 'modern slavery' (Gallagher, 2017) and 'economic migrant', which in turn affect reporting.

Moreover, UNHCR's data are not always disaggregated by age and gender, and in spite of greater efforts at multilateral cooperation, these datasets do not cover the same populations as those produced by other agencies, including the IOM. Hence, there are major differences both in coverage and categorisation of populations, which complicates analysis across datasets and undermines critical assessment of interventions.

The use of migration-specific datasets is also problematic. In the case of OECD data, foreign-born populations do not reflect

immigration status or policy categories (such as students, highly skilled migrants or refugees). Capturing such attributes is inherently difficult. First, a person's immigration status can be fluid and change quickly. For example, many international migrants who may be described as 'undocumented' or 'irregular' enter countries on valid visas and then stay in contravention of one or more visa conditions. In fact, there are many paths to irregularity, such as crossing borders without authorisation, unlawfully overstaying a visa period, working in contravention of visa conditions, being born into irregularity, or remaining after a negative decision on an asylum application has been made.

The inclusion of data from law enforcement agencies also introduces methodological challenges, not least because they tend to focus on interceptions, arrests or past activities. In the case of Frontex, one of the problems arises from its emphasis on border management. Sources of data include information on: (1) detections of illegal border-crossing; (2) detections of facilitators; (3) detections of illegal stay; (4) refusals of entry; and (5) asylum applications. This approach introduces a number of limitations, which Frontex recognises: the reliance on detection, in turn, depends on the amount of resources available (Frontex, 2015a). There is also a considerable delay before their data are published and therefore the Frontex reports offer a snapshot of previous flows and trends in migration, which may have changed by the time of publication. Further, the data are collected with a particular audience in mind, namely border-control authorities, an important but by no means exclusive audience (Frontex, 2015b). Finally, Frontex data are not immediately disaggregated by migrant category and this undermines their utility for the planning and management of migration policy beyond reception.

In spite of much standardisation among statistical reporting bodies, other data sources bring with them their own problems. While the use of censuses to collect baseline data is increasingly common, the ways in which information is recorded varies and may include self-completed postal surveys, online surveys, phone and on-the-spot interviews. For example, in England and Wales, Scotland, Northern Ireland and 13 other countries in Europe, a census is taken every 10 years where the householder receives a questionnaire in the post, completes it, and either submits it online or sends it back in the post. In other European countries, field teams visit and collect census forms from households. Annual registration is common in a number of EU countries. In Eastern Europe and in many parts of the developing world, statistical offices rely heavily on enumerators which may introduce the potential

for abuse and corruption. Moreover, some governments explicitly exclude certain groups from participating in the census; for example, Kurds in Syria were repeatedly left off the census for decades (Blitz, 2009). Their removal has complicated the correct allocation of aid and humanitarian assistance.

Public opinion datasets carry additional limitations since they assume participants have a high degree of political awareness. For example, the initial question asked by the Pew Global Attitudes Survey – 'do you approve/disapprove of the way the European Union is dealing with the refugee issue?' – assumes respondents can identify a distinctly European Union-level response. Yet, while the EU has both legal bases and polices underpinning the Common Agenda on Migration, within the EU immigration and asylum policies remain highly nationalised. Some states like the UK have opted out of the migration-specific directives and there are also many differences in terms of state practice. With the suspension of the Schengen Agreement, the closure of national borders remains an area of national – not EU -- control. These political realities may not be understood by the populations sampled.

Analysis

The above discussion illustrates the challenges of producing comparative data that accurately reflect global trends. Yet there is a growing consensus that data are essential for effective humanitarian planning and resource allocation. UNHCR (2016, p 51) has intensified its efforts to systematically collect data disaggregated by location and demographic characteristics, including in humanitarian emergencies. These data are time-sensitive, and the quicker they are collected and shared, the sooner they can be used to help those in need of assistance and to hold relevant parties to their commitments. In general, however, the quality of demographic data tends to be highest in countries where UNHCR and its partners have an operational role, undertake registration and primary data collection, and have been engaged for a year or more. Thus, while many echo the need for better data capture, in many displacement contexts national agencies cannot do this job as effectively.

While contemporary outflows from the Middle East, the Rohingya crisis, and displacement in Yemen have forced the UNHCR and its partners to improve their data collection, many problems remain. In addition to definitional issues noted above, high-intensity conflicts have made it especially difficult to operate among displaced populations in the Democratic Republic of Congo, Yemen and Syria. As a result,

the Agency and other humanitarian actors have been more reliant on local sources of information

The relevance of collecting disaggregated data in particular has become more evident as affirmed in the SDG Agenda. Even in relation to the large-scale flows across the Mediterranean in 2015-16, more disaggregated data was available in Greece, where it was directly involved in setting up the hot spots and a more effective asylum system for the country and providing accommodation, than in Italy (Kofman, 2018). Disaggregated data on vulnerable populations was fairly rudimentary during the period of mass flows halted by the implementation of the EU-Turkey deal in March 2016. After this UNHCR and the Hellenic Asylum Service began to collect data on vulnerable groups as defined by the EU Reception Directive and Greek asylum legislation.

Of the total population of concern at the end of 2015, data disaggregated by sex was available for 29.3 million – 46% of the total population of concern to UNHCR. Of the data available on sex at the end of 2015, females and males accounted for 14.3 million and 15.0 million people, respectively. Information disaggregated by age was less complete. Only 141 countries provided age-disaggregated data at the end of 2015, although coverage has increased over time. Data disaggregated by age were available for 21.2 million people (33%) at the end of 2015, compared to 17.0 million people at the end of 2014. Of the 21.2 million people covered by age-disaggregated data, 51% (10.9 million) were children.

In terms of datasets created by private organisations and NGOs, we must recognise that information is being collected at a time of great sensitivity around questions of migration and asylum and the potential for bias applies to both protectionist and pro-migrant camps. Further, the nature of current crises, which have involved many small NGOs and non-traditional actors has created additional documentary challenges that potentially impact on humanitarian responses. In the absence of archivists and information scientists who can systematically document these crises, much reporting may not be classified in a way that permits future comparisons across the many different types of data, included in NGO publications and occasional reports.

Conclusion

The above discussion suggests that the call for more and better data is fraught with challenges. These are especially relevant to those seeking to use both official and unofficial datasets for comparative research

purposes. While the UNHCR is rapidly improving its data collection, there is still a gap between humanitarian data and development-focused data produced by other UN agencies and collected through MICS, DHS, national censuses and other surveys. This is most important when considering that post-conflict situations carry additional challenges both for protecting former refugees and IDPs and gathering data on their needs.

A truly radical approach to the collection of statistical data on populations in need of protection would seek to redefine protection on the basis of demonstrable need. As we argue above, the way in which categories of concern are defined and the degree to which they are included in influential datasets has great bearing on the treatment people may receive. While all people have human rights, in practice, categories of 'concern' determine categories of perceived need and hence the allocation of resources. We suggest that additional demographic techniques are required to identify people's needs from the ground up, rather than rely on legally proscribed categories which reflect political interests and hence carry the prospect of bias and potential exclusion.

Note

[1] Refugees may be recognised on a group basis, or they may be recognised as refugees after having undergone a process of individual status determination.

References

Abel, G.J. & Sander, N. (2014) Quantifying global international migration flows. *Science* 343(6178): 1520-1522.

Bakewell, O. (1999) Can we ever rely on refugee statistics? *Radical Statistics* 72: 3-15.

Blitz, B.K. (2009) Statelessness, protection and equality. UK Department of International Development and University of Oxford Refugee Studies Centre Policy Brief, September.

Blitz, B.K. (2017) Another story: what public opinion data tell us about refugee and humanitarian policy. *Journal on Migration and Human Security* 5(2): 379-400.

Chatelard, G. (2010) What visibility conceals: re-embedding refugee migration from Iraq. In Dawn Chatty and Bill Finlayson (eds), *Dispossession and Displacement: Forced Migration in the Middle East and North Africa*. Oxford: Oxford University Press.

Edwards, A. (2013) A numbers game: counting refugees and international burden-sharing. *U. Tas. L. Rev.* **[[please give journal title in full]]** 32: 1**[[please provide full page range]]**.

Franke, M.F. (2009)**[[not cited in text]]** Refugee registration as foreclosure of the freedom to move: the virtualisation of refugees' rights within maps of international protection. *Environment and Planning D: Society and Space* 27(2): 352-369.

Frontex (2015a) *Eastern European Borders Annual Risk Analysis 2015*. Warsaw: Frontex, May.

Frontex (2015b) *FRAN Quarterly – Quarter 4 October–December 2014*. Warsaw: Frontex, April.

Gallagher, A.T. (2017) What's wrong with the Global Slavery Index? *Anti-Trafficking Review* 8: 90-112.

Gemenne, F. (2011)**[[not cited in text]]** Why the numbers don't add up: a review of estimates and predictions of people displaced by environmental changes. *Global Environmental Change* 21: S41-S49.

Harrell-Bond, B., Voutira, E. and Leopold, M. (1992) Counting the refugees: gifts, givers, patrons and clients. *Journal of Refugee Studies* 5(3-4): 205-225.

Kofman, E. (2018) Gendered mobilities and vulnerabilities: refugees crossing to and through Europe. *Journal of Ethnic and Migration Studies*'**[[please provide volume and page numbers]]**

IOM and McKinsey (2018) *More than Numbers: How Migration Data can Deliver Real-Life Benefits for Migrants and Governments*. Geneva: International Organization for Migration and McKinsey.

Laczko, F., Singleton, A., Brian, T. and Rango, M. (2016)**[[not cited in text]]** Migrant arrivals and deaths in the Mediterranean: what do the data really tell us? *Forced Migration Review* (51): 30.**[[please provide volume and page numbers]]**

Merry, S. (2017)**[[not cited in text]]** Counting the uncountable: constructing trafficking through measurement. In P. Kotiswaran (ed), *Revisiting the Law and Governance of Trafficking, Forced Labor and Modern Slavery* (Cambridge Studies in Law and Society, pp. 273-304). Cambridge: Cambridge University Press.

O'Connell Davidson, J. (2014)**[[not cited in text]]** The making of modern slavery: whose interests are served by the new abolitionism? *British Academy Review* 24 (Summer): 28-31.

UN High Commissioner for Refugees (UNHCR) (2010) *UNHCR and De Facto Statelessness*, April, LPPR/2010/01.**[[please provide publisher location and name]]**

UN High Commissioner for Refugees (UNHCR) (2017)**[[2016 in text?]]** *Global Trends. Forced Displacement in 2016*. UNHCR.**[[please provide publisher location]]**

Zetter, R (2007). More labels, fewer refugees: remaking the refugee label in an era of globalization. *Journal of Refugee Studies* 20(2): 172–192.