Chapter 1

Sources, methods and data considerations

Introduction

The *1951 Convention relating to the Status of Refugees*, and other relevant international instruments, give UNHCR a responsibility to compile and to provide reliable statistics on the number and condition of refugees and other populations of concern to the organization. The Office also collects and disseminates information on IDPs, including statistics on their number and condition. UNHCR monitors returnees (both refugees and IDPs) to determine the magnitude and sustainability of returns. As a result, UNHCR engages with governments and partners within the humanitarian community to not only provide a comprehensive picture of the number but also of the protection gaps and needs of refugees, IDPs, stateless persons, returnees and other populations of concern to the Office. This task is complex due to factors influencing the availability, timeliness, quality and comparability of the statistics, thus complicating the provision of a comprehensive picture of the population of concern to UNHCR.

In addition to definitional challenges, the changes in the numbers of those of concern to the Office can be explained by a variety of reasons, including demographic factors (e.g. individual or mass population movements); legal increases and decreases (e.g. cessation clause or recognition rates); and other changes in the reporting of the statistics. Second, some refugees have not been individually registered because of a lack of awareness of the need to register or the perception that the risks of registration are higher than its benefits. For example, individuals in mixed migration flows who would potentially be deemed by UNHCR to be in need of international protection may refrain from seeking asylum in a given country due to fear of being deported together with other migrants. Third, an increasing number of countries grant complementary or subsidiary forms of protection, which are more difficult to compare internationally. For instance, people fleeing conflict from the same location may be recognized as refugees by some countries, but not by others. Fourth, administrative records may overreport the actual number of people due to registration errors. This is particularly the case when registration is linked to the provision of services or benefits. Fifth, a range of data collection methods and sources are used by different countries or for different categories of persons of concern, making the provision of comparable statistics a difficult task.

In order better to comprehend the observed trends, understanding the complexity of UNHCR data is critical. The main purpose of this chapter is to present information that will assist in interpreting UNHCR's data by addressing some of the constraints raised above. This information, also known as 'metadata', provides an understanding of the basic concepts behind the data in the 2008 Statistical Yearbook, including definitions, classifications, the underlying estimation methods and key aspects of data quality. Most of the statistics reported in this Yearbook have been collected through the Annual Statistical Report[11] and generally reflect the definition and data collection methods of governments, except from those countries where only UNHCR registration systems are used.

---

11   The Annual Statistical Report is the official data collection form completed by all UNHCR country offices.

Definitions and scope

UNHCR identifies seven population categories, collectively referred to as "persons of concern". This includes: (a) refugees; (b) asylum-seekers; (c) internally displaced persons (IDPs); (d) refugees who have returned home (returnees); (e) IDPs who have returned home; (f) stateless persons; and (g) other people who do not fall under any of the above categories but to whom the Office extends its protection and/or assistance activities. In 2007, two additional sub-categories were introduced: (1) people in refugee-like situations (included under refugees); and (2) people in IDP-like situations (included under IDPs).

Refugees include individuals recognized under the *1951 Convention relating to the Status of Refugees*; its *1967 Protocol*; the *1969 OAU Convention Governing the Specific Aspects of Refugee Problems in Africa*; those recognized in accordance with the UNHCR Statute; individuals granted complementary forms of protection[12]; or, those enjoying "temporary protection"[13].

The refugee population also includes people in a refugee-like situation, most of whom were previously included in the Others of Concern group. This sub-category is descriptive in nature and includes groups of people who are outside their country or territory of origin and who face protection risks similar to refugees, but for whom refugee status has, for practical or other reasons, not been ascertained.

Asylum-seekers ('pending cases') are individuals who have sought international protection and whose claims for refugee status have not yet been determined. Those covered in the Yearbook refer to claimants whose individual applications were pending at the end of 2008, irrespective of when they may have been lodged.

Internally displaced persons are people or groups of individuals who have been forced to leave their homes or places of habitual residence, in particular as a result of, or in order to avoid the effects of armed conflict, situations of generalized violence, violations of human rights or natural- or human-made disasters, and who have not crossed an international border.[14]

As part of the cluster approach[15], UNHCR is involved in the collective response to situations of internal displacement in selected countries. For purposes of UNHCR's statistics, however, this population only includes conflict-generated IDPs to whom the Office extends protection and/or assistance. Therefore, UNHCR statistics do not provide a comprehensive picture of global internal displacement.[16]Moreover, UNHCR's IDP statistics are not necessarily representative of the entire IDP population in a given country, as they are limited to those who are protected and/or assisted by the Office. While UNHCR, within the cluster approach, provides support to both IDPs and other affected persons, the latter category is not included in the statistics.

---

12   Complementary protection refers to protection provided under national or regional law in countries which do not grant 1951 Convention refugee status to people who are in need of international protection against serious, but indiscriminate risks.

13   Temporary protection refers to arrangements developed by States to offer protection of a temporary nature to people arriving from situations of conflict or generalized violence without the necessity for formal or individual status determination. This usually applies to situations of large-scale influx.

14   Source: Handbook for the Protection of Internally Displaced Persons (provisional release), p.6, Global Protection Cluster Working Group, December 2007 (www.humanitarianreform.org).

15   In December 2005, the Inter-Agency Standing Committee endorsed the Cluster Approach for situations of internal displacement. Under this arrangement, UNHCR assumes leadership responsibility and accountability for three of the nine "clusters", namely: protection; emergency shelter; and camp coordination and camp management.

16   For global IDP estimates, consult the Internal Displacement Monitoring Centre (IDMC) of the Norwegian Refugee Council (NRC) website at www.internal-displacement.org.

The 2008 IDP population category includes people in IDP-like situations. This sub-category is descriptive in nature and includes groups of people who are inside their country of nationality or habitual residence and who face protection risks similar to IDPs but who, for practical or other reasons, could not be reported as such.

Returned refugees (returnees) refer to refugees who have returned voluntarily to their country of origin or habitual residence. For purposes of this Yearbook, only refugees who returned between January and December 2008 are included. However, in practice, operations may assist returnees for longer periods.

Returned IDPs refer to those IDPs who were beneficiaries of UNHCR's protection and assistance activities and who returned to their areas of origin or habitual residence between January and December 2008. However, in practice, operations may assist IDP returnees for longer periods.

Stateless persons are individuals not considered as nationals by any State under national laws. UNHCR statistics also include *de facto* stateless persons and persons with undertermined nationality. UNHCR has been given a global mandate by the United Nations General Assembly to contribute to the prevention and reduction of statelessness and the protection of stateless persons. UNHCR's Executive Committee has requested the Office to report regularly on the magnitude of the phenomenon. The Office also has specific functions under Article 11 of the *1961 Convention on the Reduction of Statelessness* to receive claims from persons who may benefit from the safeguards contained in that Convention and to assist them and the States concerned to resolve those claims.

Statelessness may be an important reason for fleeing one's place of origin. However, most stateless persons have not fled their country of habitual residence. Where stateless persons are also refugees, they are reported under UNHCR's refugee statistics. Because stateless persons are often difficult to identify, UNHCR does not have comprehensive country-level data but only a rough estimate of the total magnitude of statelessness worldwide (see also Chapter 2).

Other groups or people of concern refers to individuals who do not necessarily fall directly into any of the groups above but to whom UNHCR has extended its protection and/or assistance services, based on humanitarian or other special grounds.

Data sources and providers

States bear the main responsibility to provide data on displaced populations to UNHCR. Nevertheless, the number of data sources used to establish the size and characteristics of the population of concern to UNHCR varies within and among countries. Governmental agencies and UNHCR offices are the main data providers, either separately or jointly. NGOs are also key providers of refugee and IDP data in a number of countries. These three sources account for more than 90 per cent of UNHCR's refugee statistics. Registers, surveys, estimation, registration processes or censuses are the main methods and tools used to determine the figures. Although one source may provide all the data within a country, more commonly more than one source to obtain data on UNHCR's population of concern is used.

In 2008, information on data sources was provided by some 156 countries, with a number reporting more than one source. The host government was the sole data provider in 55 of these countries (35%). This represents a 6 per cent increase compared to 2006 (45 countries) and 8 per cent against 2004 (41 countries). This development is a reflection of the increasing capacity of host governments to register and enumerate refugees and others of concern.

Nevertheless, in a significant number of countries, the capacity to provide refugee statistics is still limited and this responsibility remains with UNHCR. By the end of 2008, UNHCR was reported as the single source of refugee statistics in 54 countries (35%) while in 23 countries (15%) the collection and provision of statistics was a combined undertaking of the government and UNHCR. For the latter, a continued decrease has been observed since 2006 when one out of four countries (25%) reported a joint responsibility.

Multiple sources, including in some instances NGOs or UNHCR and States, in addition to others, were reported in 24 countries (15%). In total, governments and UNHCR have been quoted as the single source or one of the main sources in over 130 countries.

Fig I.1 Sources of refugee data, 2008
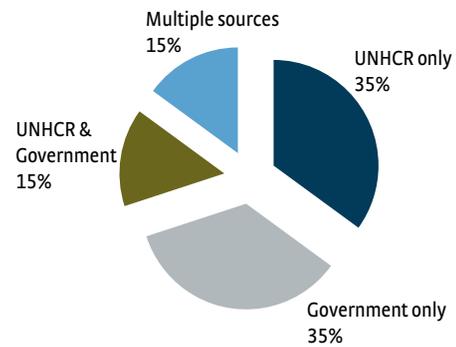
Data collection methods

UNHCR, together with its partners, uses a wide range of data collection methods, including registration/registers, surveys, censuses, and estimations. Each method has its own strengths and limitations. To make data as comparable as possible, UNHCR ensures that statistics from all these sources are triangulated. This requires regular interaction with and feedback to the various stakeholders who collect and compile refugee statistics, especially UNHCR offices, local NGOs and government officials.

Registration data is a vital statistical source for planning and management purposes. Refugees can only be adequately protected when they are registered. Refugee statistics are generally based on individual registration records, kept by the host government. As such, the registration of refugees is a government responsibility. But in cases where the government is unable to register persons of concern, UNHCR may assist or undertake registration activities. So far, UNHCR's registration software *proGres* has been rolled-out in 72 countries, including five in which it has been adopted by the government. During the registration process, data is collected, verified and updated. Such data includes name, sex, date of birth, country of origin, relationship within household, marital status, place of displacement and specific needs. This information is normally stored in a database and should be continuously updated to reflect changes in the population. As

Fig I.2 Trends in sources of refugee data (%)
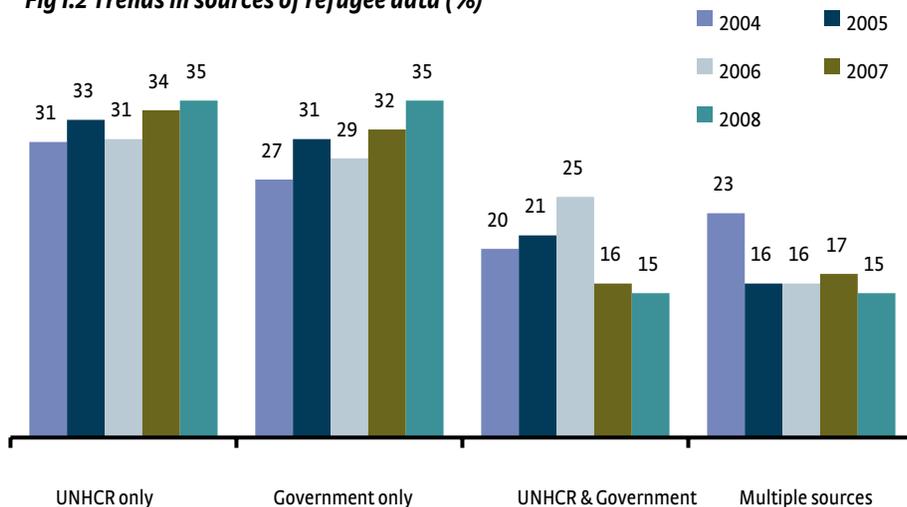
pressures to register populations are greater than those to de-register them, data may become inflated, requiring periodic verification. Registers may exclude parts of the population, as not all those who qualify are able or willing to register themselves. Periodic registration exercises are one of the main reasons for a revision of the population estimates. By the end of 2008, registration was the only method used in 95 countries (60%), while in 16 other countries (10%) it was combined with estimation methods.

Estimation is the second main data collection method used by UNHCR and its partners. By using correlated indicators on existing data, it is possible to obtain a measure of the value of an unknown quantity in a population or a sample. This method is primarily used in two contexts. First, considering the operational and legal difficulties in registering IDPs on an individual basis and following the recommendation of the Global Protection Cluster Working Group[17], most humanitarian agencies rely on *IDP profiling* (see also below on Surveys). This allows them to obtain a better picture of IDP populations and to answer questions about them such as where they are located, what are the difficulties of distinguishing IDPs from host communities, and how to compile baseline information for protection, programming and advocacy purposes.

Fig I.3 Basis of refugee data, 2008
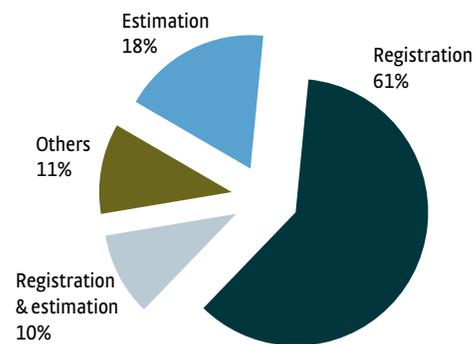
Second, as many industrialized countries do not maintain a dedicated refugee register and are thus not in a position to provide accurate information on the number of refugees residing on their territory, UNHCR estimated the refugee population in 24 countries based on the recognition of asylum-seekers over a 10-year period. This cut-off period is assumed to reflect the average time it generally takes for a refugee to acquire the nationality of his/her host country.

Estimation as the sole source is used in 29 countries (18%), most of which are industrialized. This method is combined with registration in 16 other countries (10%).[18] Below is an example of estimation:

Example: Italy
– Asylum-seekers granted Convention refugee status, 1999-2008: 12,448
– Asylum-seekers granted a complementary form of protection and/or allowed to remain for humanitarian reasons, 1999-2008: 34,613
– UNHCR estimate of the refugee population in Italy, end of 2008: 47,061

The "Other" category in Figure I.3 includes mainly surveys and censuses, either alone or in combination with registration and estimation. In some countries, more than two methods are used, making data comparability difficult. Surveys and censuses are excellent sources for planning purposes, but do not easily allow follow-up at the individual level. Although a census includes information on each individual, it provides only a "snapshot" of the situation which quickly becomes outdated. Surveys are cost-effective, but unless carried out according to strict rules, they may not provide a representative picture. In general, various data methods are combined and triangulated to obtain the best possible result. Similar to 2007, by the end of 2008, 18 countries (11%) obtained refugee figures through surveys, censuses and/or combined methods.

---

17   See www.humanitarianreform.org

18   Because of a significant number of countries having access to *proGres*, estimation is rarely used in non-industrialized countries.

Key characteristics presented in the 2008 Yearbook

In addition to presenting the numbers and trends of those of concern to UNHCR, the 2008 Statistical Yearbook contains a variety of population characteristics, broken down according to demography (mainly sex and age) and location (type of settlement such as camp or urban setting). This type of information, however, is only partially available for persons of concern to UNHCR (62 per cent for sex and 42 per cent for age).

Location refers to the type of settlement of populations: camp, urban or rural/dispersed areas. UNHCR has location data available for 70 per cent of all refugees under its mandate. When including the other population categories (IDPs, returnees, etc.), this figure drops to 43 per cent. While much progress has been made in collecting information on basic factors such as age, sex and location, much more needs to be done to improve and expand the understanding of these populations.

The Yearbook also provides information about the type of recognition (individually recognized vs. group based/*prima facie*) under which refugee status was granted. With regards to asylum-seekers, information about the level and type of procedures as well as recognition rates is also included. Selected indicators on food, nutrition and protection also feature in this Yearbook.

A detailed description and explanation of the above and other characteristics will be presented in the corresponding chapters.

Other data considerations

Statistics concerning the former USSR have been reported under the Russian Federation, while those concerning the former Czechoslovakia have been reported under the Czech Republic. Data concerning the former Yugoslavia as well as Serbia and Montenegro have been reported under Serbia. In most Annex tables, the word "Democratic" has been abbreviated to "Dem.", whereas the word "Republic" is often reflected as "Rep." The former Yugoslav Republic of Macedonia is referred to as TfYR Macedonia.

Figures below 1,000 are generally rounded to the nearest 10; figures between 1,000 and 10,000 are rounded to the nearest 100; figures between 10,000 and 100,000 are rounded to the nearest 1,000; figures between 100,000 and 1 million are rounded to the nearest 10,000; and figures above 1 million are rounded to the nearest 100,000.

Unless specified otherwise, the 2008 Statistical Yearbook does not refer to events occurring after 31 December 2008.

Box 1.
From survey to protection: the case of stateless persons in Kyrgyzstan

Since gaining independence in 1991, Kyrgyzstan has taken positive steps to reduce and prevent statelessness. The country did not, however, establish a proactive mechanism to identify, determine, register and document stateless persons and did not collect any statistics on the number of stateless persons residing on its territory. Estimates thus varied between several hundred and hundreds of thousands of affected persons.

Until 2008, UNHCR utilized the first national census of 1999 as the only available baseline data, regularly decreasing the number of persons who had identified themselves as being stateless in this census by those having been granted Kyrgyz citizenship. This method was unsatisfactory as it did not take into account the unreliability of the original data source, which was based on self-identification, population movements, or the creation of new cases of statelessness. Moreover, it did not provide a proper age or sex breakdown of stateless populations.

Upon request of the Government, in 2007-2008, UNHCR commissioned expert NGOs to conduct three surveys in the country. The goal of two studies in the north and the one in the south was to estimate more accurately the current total number and situation of affected persons and to investigate in more detail the reasons for them being stateless.

With the help of village heads and local passport departments, field teams interviewed thousands of individuals without proper documentation in districts with high levels of population movement, and gathered secondary data on other areas. Altogether, the surveys identified close to 13,000 individuals who were stateless, or at risk of becoming so.

The surveys gathered data disaggregated by sex, age, diversity and main reasons for statelessness. It showed that more than 70 per cent of "former citizens of another state, who have become stateless because of shortcomings in citizenship legislation and practices" are female, and that the majority reside in districts with a high level of mobility. Another 4,500 respondents still held old USSR passports and had for various reasons failed to exchange them with Kyrgyz ones even though most of them had been living in the country already for many years. The surveys also identified more than 6,000 children of stateless persons who faced difficulties in obtaining Kyrgyz nationality due to the status of one or both of their parents.

The surveys noted regional differences in the ethnic composition of the stateless population in this multi-ethnic country: half of respondents in the south were ethnic Kyrgyz and half of a different ethnic background, while in the north these figures were 71 and 18 per cent, respectively. The ethnic composition of stateless persons also varied between provinces: in Batken Province, in the ultimate south of Kyrgyzstan, for example, 90 per cent of stateless persons were ethnic Kyrgyz.

A legal analysis which formed part of the survey in the south confirmed that statelessness in Kyrgyzstan was mainly caused by incomplete implementation of otherwise quite advanced legislation. This problem was addressed in September 2009, when UNHCR and the Presidential Administration co-hosted a conference on statelessness. During this conference, parliamentarians, government officials, judges and NGOs discussed and adopted a National Action Plan to Prevent and Reduce Statelessness in the Kyrgyz Republic, drafted by an expert working group using the surveys' findings as a basis.

The first phase of implementing this Action Plan, during the last quarter of 2009, focused on holders of USSR passports and on the drafting of necessary by-laws and instructions to improve the implementation of the national citizenship law. A further study and recommendations of the Conference concluded that Kyrgyzstan should accede to the two UN Statelessness Conventions, requiring only minor amendments in national legislation. Full implementation of the Action Plan in 2010 and 2011 will greatly reduce and prevent statelessness in the Kyrgyz Republic.

The Action Plan also plans to create administrative procedures to pro-actively identify, register and document stateless persons. Once established, these mechanisms should allow for the keeping and updating of adequate official statistics.

The impact of the surveys commissioned by UNHCR was multiple. They proved to be an invaluable tool to gain insights into a problem whose dimension had been largely unknown. They, for the first time, provided reliable and disaggregated data on the problem of statelessness. They served as a basis for structured discussions leading to concrete actions that solved the problems identified. Last but not least, they helped put in place sustainable mechanisms for future data gathering which should make more on-the-spot surveys superfluous.