United Nations E/CN.9/2016/3

Economic and Social Council

Distr.: General
27 January 2016
Advance unedited version

Original: English

Commission on Population and Development
Forty-ninth session
11-15 April 2016
Item 3 of the provisional agenda*
Actions for the further implementation of the Programme of Action of the International Conference on Population and Development

Strengthening the demographic evidence base for the post-2015 development agenda

Report of the Secretary-General

Summary

In its decision 2015/101, the Commission on Population and Development established that the special theme for its forty-ninth session would be "Strengthening the demographic evidence base for the post-2015 development agenda".

In September 2015, the General Assembly adopted the 2030 Agenda for Sustainable Development, which includes 17 ambitious goals to ensure a sustainable future for humanity. The sustainable development goals (SDGs), together with 169 targets that further specify the achievements expected, will guide the actions of Governments and the development community through 2030. Monitoring progress toward the achievement of the goals and targets will be crucial to guide action. Reliable and timely demographic information is essential for effective monitoring.

This report reviews the current status of demographic evidence and its capacity to support SDG monitoring. It focuses on the core systems generating demographic data and discusses how the "data revolution" has already improved their performance and the potential for accelerating such improvements.

In addition, the report underscores that open access to anonymized micro-data and routine use of geo-referencing in data collection are necessary to expand the use of demographic data. Geo-referencing makes possible the integration of diverse datasets and facilitates disaggregation and the incorporation of big data in analyses and modelling.

The report also highlights the role of institutions producing consistent and comparable sets of demographic indicators essential for monitoring the SDGs globally.

This report was prepared by the Population Division of the Department of Economic and Social Affairs.

* E/CN.9/2016/1.

I. Introduction

1. The 2030 Agenda for Sustainable Development, adopted by the General Assembly in September 2015, provides an ambitious set of goals whose achievement will set humanity on a path towards sustainability. Progress made toward achieving the Agenda's 17 sustainable development goals (SDGs) and 169 targets will need to be monitored closely to provide timely guidance for effective action. The data requirements partially overlap with, and in some cases go beyond those needed to monitor the Programme of Action of the International Conference on Population and Development (ICPD)[1] and the regional commitments for the implementation of the Programme of Action.

2. The process of establishing indicators to measure progress in the SDGs is well under way. Timely and reliable demographic data will therefore be necessary both to calculate indicators and to provide weighting factors to obtain regional or global estimates. Furthermore, the emphasis of the 2030 Agenda on inclusiveness and the need to leave no one behind implies that demographic data, in particular, will have to be available for specific population groups, such as children, women of reproductive age, youth or the elderly, and by location, including for small areas within countries.

3. The increased demands that monitoring the SDGs will put on national statistical systems and other institutions generating relevant information have been the subject of increased attention. The numerous initiatives to improve crucial statistical systems , and exploiting advances and innovation in information and communication technology (ICT) to improve the evidence-base, particularly in developing countries, has been the focus of the Secretary-General's Independent Advisory Group on a Data Revolution for Sustainable Development[2] and the Partnership in Statistics for Development in the 21st Century (PARIS21). Furthermore, in its *Data for Development Action Plan to Finance the Data Revolution for Sustainable Development,* the UN Sustainable Development Solutions Network identifies eight data "instruments" as necessary for monitoring the SDGs, three of which—censuses, civil registration and vital statistics (CRVS), and household surveys—are core instruments for the generation of demographic information.

[1] Including chapter XII, covering data and research requirements, and recommended actions.
[2] *A World That Counts*, http://www.undatarevolution.org/report/#nav-mobile, accessed on 20 November 2015.

4. Improving the reliability, timeliness and accessibility of demographic data needs to be a central focus of any effort to strengthen statistical systems for monitoring the SDGs. This report reviews the current status of demographic data and suggests strategies to improve the core instruments used to generate them. It focuses mainly on data permitting the estimation of population dynamics—population counts, births, deaths and migrants—and on data about the proximate determinants of fertility.

5. Improving data-generation systems, however, is only part of a comprehensive strategy to improve the demographic evidence-base. It is also crucial to foster and strengthen initiatives to compile, harmonize, consolidate, disseminate and analyse the data. Open data policies that allow researchers and decision-makers to access micro-data[3] with safeguards to ensure confidentiality are essential to promote in-depth analysis and utilization of rich datasets. Costly data-generation projects often have little impact because the data produced remain inaccessible to analysts.

6. A major innovation that should become routine practice in data generation is geo-referencing. The integration of diverse data sources via geo-referencing can produce vastly enriched evidence linking demography with other relevant processes. For instance, studying the effects of climate change on people, requires access to integrated data on population and the environment. Programme planning, implementation and service delivery can also greatly benefit from geo-referenced information, especially to monitor and guide the reduction of subnational inequities in service delivery and health status. Such integrated data are necessary to guide action promoting the wellbeing of both people and the planet.

II. The core sources of demographic data

A. Population censuses

7. In most countries, population and housing censuses are the only source of overall population numbers by age and sex, marital status, educational attainment, occupation, ethnicity, migrant status, household composition, housing characteristics

[3] "Micro-data" refers to the anonymized dataset(s) containing the information collected about each observation unit (e.g. housing unit, household or person).

.

and other relevant socio-demographic characteristics. Because censuses aim to enumerate all people in a country at a particular time, they yield data for small groups and small areas that other data sources cannot produce. Censuses are therefore essential to assess whether any group is being left behind in the development process.

8. Most countries[4] conduct a population census at least once every decade. During 2005-2014, there were censuses to count the populations of 214 countries or areas, covering 93 per cent of the world's population.[5] Only 20 did not: 8 in Africa, 8 in Asia, 2 in Europe and 2 in Latin America and the Caribbean. This represents an improvement with respect to the prior rounds of 1995-2004 and in 1985-1994, in which censuses were conducted for only 206 countries or areas. Nevertheless, in preparing the 2015 Revision of population estimates, the UN Population Division found that for 17 per cent of countries in Africa, 8 per cent of those in Asia and Latin America and the Caribbean, and 4 per cent of those in Oceania, the most recent population counts available referred to years prior to 2005.

9. Population censuses are major operations conducted at lengthy intervals. Censuses typically start by updating the cartography of a country's inhabited areas and thus provide an excellent means of geo-referencing data on the location of people and households. However, censuses are not meant to monitor change continuously. Innovative approaches to gather census-type data at shorter intervals are being used by some countries. Multi-modal approaches rely on continuous or ad-hoc surveys to supplement census and registers. For example, the Netherlands conducted a virtual census by integrating data from registers and multi-purpose surveys, while the United States of America has used the annual American Community Survey to update the results of its 2010 census. France has been conducting a "rolling census" since 2005, consisting of repeated surveys that, over time, cover the whole population. Countries of the Gulf Cooperation Council are considering options to conduct register-based censuses.

10. ICT innovations can improve every stage of census operations including cartographic updates, logistics monitoring, questionnaire design, and data collection, coding, storage, analysis and dissemination. However, countries vary widely in their

[4] In this report, the term "countries" includes areas lacking an independent political status and that are generally a portion of one or more independent States.

[5] http://unstats.un.org/unsd/demographic/sources/census/2010_PHC/censusclockmore.htm

use of ICT. Most countries still rely on face-to-face interviews and paper questionnaires to gather census data, but scanning and imaging technology is commonly employed to digitize information. Computer editing and processing are virtually universal and the use of the geographic positioning system (GPS) to produce geo-referenced data has expanded.

11. The use of ICT to disseminate and provide access to census information is far from universal. Responses from 121 national statistical offices (NSOs) that had conducted a census by 2011, showed that just 28 per cent of them would disseminate results via static web pages and even fewer (14 per cent) would do so through interactive online databases. Paper publications were the main dissemination medium for over half of the NSOs responding.[6] When census information is disseminated only via paper publications, the type and amount of information accessible to users is limited. A positive development in this regard is the growing number of national statistical offices that have been disseminating selected national and subnational summary indicators through interactive desktop or online platforms, such as CensusInfo, supported by UNSD, UNICEF and UNFPA, and PC-AXIS, supported by Statistics Sweden.

12. But to utilize fully the rich data produced by censuses, availability of the micro-data is required. Software designed specifically to make micro-data accessible via online interactive databases with safeguards to maintain confidentiality should be used to disseminate census results. To that end, the UN Statistics Division, in partnership with UNICEF and UNFPA, has developed CensusInfo, software to disseminate census results at any relevant geographical level.[7] CELADE, the Population Division of ECLAC, has developed REDATAM, software that allows the processing and visualization of census information and the development of applications for analysis. REDATAM has been used by at least 32 countries to disseminate census micro-data online, 24 of which provide access to census information collected since 2005. Those 32 countries include 17 in Latin America, 6 in the Caribbean, 7 in Africa and 2 each in Asia and Oceania.[8] The IPUMS-International “Analyze Data Online” system currently facilitates statistical analysis

[6] US Census Bureau, Mid-Decade Assessment of the United Nations 2010 World Population and Housing Census Programme, 2013, ESA/STAT/AC.277/1

[7] http://www.censusinfo.net/

[8] http://www.cepal.org/cgi-bin/getprod.asp?xml=/redatam/noticias/paginas/5/14185/P14185.xml&xsl=/redatam/tpl/p18f.xsl&base=/redatam/tpl/top-bottom.xsl

of integrated, high-precision census samples for 80 countries in roughly equal proportions by continent.[9] The number of countries and censuses included in this system has been increasing over time.

13. Although NSOs report that use of ICT improves data quality and timeliness, they also acknowledge that major barriers to adopting ICT innovations are a lack of trained personnel and cost. Flexibility and a more widespread adoption of technological innovations will be essential to improve census taking in the 2020 round. Data collection can be enhanced by using hand-held devices for direct data input, thereby reducing data-capture costs. GPS, GIS and new mapping techniques will likely improve address listing and geo-referencing. Data dissemination using online databases and specialized software will expand the user-base and allow a deeper exploitation of the data. These developments should be accompanied by adopting appropriate open-data policies and adequate funding and training of the required personnel in the use of new technologies.

14. At the international level, the Statistical Commission sets standards and provides guidance for conducting censuses. In preparation for the 2020 round, the Commission adopted in 2015 the third revision of *Principles and Recommendations for Population and Housing Censuses*,[10] which makes recommendations on census content, presents alternative methodologies for census-taking, and discusses technologies to improve census operations. At the regional level, the regional commissions are in charge of complementing global recommendations by providing regional guidance.

B. Population Registers

15. A population register is a system that records on a continuous basis selected information about each resident of a country. Both the organization and the operation of a population register have a legal basis. While the main purpose of population registration is administrative, the register can be the source of up-to-date statistics on the size, characteristics and location of the population and of its components of change, namely, births, deaths and changes of residence, which allow the estimation of internal and international migration. When population registers are

[9] https://international.ipums.org/international/sda.shtml
[10] http://unstats.un.org/unsd/demographic/meetings/egm/NewYork/2014/P&R_Revision3.pdf

centralized, that is, when the information they produce for a whole country is maintained together, they permit to ascertain the status of a population and its socio-demographic structure at any time and for any location, including small areas. Central registers also permit studying families and their living arrangements and performing longitudinal analysis of demographic events in the lives of individuals. For these reasons, well-maintained, computerized population registers represent the best source of demographic data.

16. However, few countries maintain population registers. They are most common in Europe. At least 23 European countries have population registers and in 20 of them the registers are centralized or in the process of being centralized.[11] In Asia, at least seven countries have some form of population registration but do not use the registers to generate demographic statistics.

17. A number of European countries with registers do not conduct conventional population censuses. Austria, Denmark, Finland, Iceland, Norway and Sweden produce census-like data entirely from linked administrative databases that include the central population register. Belgium, Germany, the Netherlands and Slovenia, derive population data from population registers and use surveys to gather supplementary information. The Czech Republic, Estonia, Italy, Latvia, Lithuania, Poland, Spain and Switzerland conduct censuses with preprinted information derived from the register to update such information.

18. One shortcoming of population registers is they do not always record all changes of residence, particularly those of foreigners that leave the country, which can affect significantly the accuracy of population counts.

19. Nonetheless, given the vast potential that population registers have to improve national demographic statistics, countries that already have a population registration system should consider using the register to generate demographic statistics. In those countries, evaluating the operation of the existing register and developing a strategy to correct any deficiencies detected would be a first step in making it a viable statistical source. If deficiencies in coverage are detected, such as registering departures and undocumented migrants, a campaign to inform people about the benefits of registering and to prompt them to act will be required. India provides an interesting example: it is in the process of establishing a state-of-the-art registration

[11] Michel Poulain and Anne Herm, “Central population registers as a source of demographic statistics in Europe”, Population-E, 2013, vol. 68, No. 2, pp. 183-212.

system where every resident is issued an identity card with a unique identification number and embedded biometric information. A major campaign to inform and persuade the population to register has been key to its success.

20. Many countries maintain administrative registers to record the entry, residence and employment of the foreign population. Given the paucity of data on international migration, the tabulation and publication of data on applications for visa, residence authorizations and labour permits in critically important.

B. Civil registration and vital statistics

21. Civil registration is an administrative system that records vital events—births, deaths, marriages and divorces— as they occur. It issues legal documents that establish the legal identity and family relationships of individuals. To be effective, civil registration has to be compulsory, universal, continuous, permanent and confidential. The registration records should be archived and maintained so that they can be retrieved when needed. Coordination among civil registrars, health and statistical authorities is needed to ensure the timely production of vital statistics.

22. Although complete vital statistics are essential for demographic accounting, many countries lack them. Among 234 countries or areas, only 60 per cent register at least 90 per cent of births and just 56 per cent register at least 90 per cent of deaths. Furthermore, only 39 per cent have reasonably complete data on causes of death and 41 per cent lack data on most causes. In Africa, 71 per cent of countries fail to report vital statistics. Globally, between 2005-2009 and 2010-2014 the number of countries failing to report vital statistics to the UN Statistics Division increased from 27 per cent to 33 per cent for births and from 26 per cent to 34 per cent for deaths. The absence of timely reporting impairs global monitoring of vital statistics

23. Yet, DHS and MICS surveys show that between 2004 and 2014, the percentage of children under five who were registered increased from 58 to 65 per cent. Despite this improvement, in 2015, 230 million children under five lacked a birth certificate, including 85 million in Africa and 135 million in Asia and Oceania. The SDGs call for the registration of all births so that every person has a legal identity. Governments should ensure that the capacity to produce vital statistics expands in parallel to the expansion of birth registration.

24. The UN Convention on the Rights of the Child, which entered into force in 1990, establishes that "the child shall be registered immediately after birth" (Article 7). UNICEF has therefore been fostering the expansion of birth registration for over two decades. However, only recently has the urgent need for reliable data on the mortality of mothers and children focused attention on improving CRVS systems. In 2011, the UN Commission on Information and Accountability for Women's and Children's Health (COIA) called on Governments to establish CRVS systems, recognizing CRVS as the most efficient source of data to monitor child and maternal mortality.[12] With COIA's support, 51 of the 75 priority countries have completed assessments of their CRVS systems and 28 of them have developed multi-sectoral plans of action.

25. In parallel, development partners are aligning forces to support CRVS development in priority countries. The World Bank estimates that over a decade US$3.82 billion will be necessary to establish or improve CRVS systems in 73 priority countries (China and India are excluded from this estimate). [13] Most of these investments are expected to come from national budgets but donors and development partners would have to support the early stages of scale-up and cover a financing gap estimated at US$199 million per year over a decade.

26. The Government of Canada has already prioritized funding for CRVS within its US$3.5 billion phase I Muskoka commitment.[14] In March 2015, Bloomberg Philanthropies in collaboration with the Australian Department of Foreign Affairs and Trade established a global partnership that will bring together key expertise in the development of CRVS to improve vital statistics in 20 developing countries. At the World Bank, the Global Financing Facility for Reproductive, Maternal, Newborn, Child, and Adolescent Health (GFF) will provide results-focused financing while supporting the systems needed to monitor progress and measure results, in particular CRVS.

27. At the regional level, meetings at the ministerial level have committed to strengthen CRVS, while the regional commissions are providing technical support, including for the implementation of action plans adopted by ECA and ESCAP. Their task should be guided by the third revision of *Principles and Recommendations for a*

---

12 Commission on Information and Accountability for Women's and Children's Health, *Keeping Promises, Measuring Result*s.: *Final Report*, http://www.who.int/topics/millennium_development_goals/accountability_commission/Commission_Report_advance_copy.pdf.

13 World Bank, *Global Civil Registration and Vital Statistics: Scaling–up Investment Plan 2014-2025,* 2014.

14 http://mnch.international.gc.ca/en/topics/leadership-ongoing.html

Vital Statistics System[15] approved by the Statistical Commission in 2014. Furthermore, WHO has been promoting innovative approaches to improve mortality statistics within CRVS[16] and has developed technical guidance on the collection, analysis and use of mortality statistics, focusing on the implementation of the International Classification of Diseases (ICD) and on the use of verbal autopsy techniques to generate information on causes of death at the community level.[17]

28. These initiatives are long overdue. Only a third of deaths worldwide are registered and, in some countries, nearly half are attributed to generic causes that are uninformative for public policy.[18] Although over the past decade some CRVS systems have improved, others have deteriorated or ceased producing data. Periodic assessments of the performance of the CRVS system, including at the sub-national level, should be standard practice.

29. Improving CRVS systems goes beyond technical issues. Clarity about the responsibilities of the different government units involved and a streamlined process leading from registration to statistics are equally important. Close collaboration between CRVS authorities and the health system is essential. Registration is most efficient when it is done where births and deaths occur, including in clinics or hospitals. Burial services should be required to request death certificates. Registration fees or requirements for prior documentation from those registering an event are disincentives that should be eliminated or minimized.

30. A concerted and serious effort to expand and improve CRVS systems, particularly in low-income and lower-middle-income countries, is a priority. Few would doubt that reducing premature death from chronic and degenerative diseases is desirable. Yet, precisely those countries where the prevalence of such diseases is expected to increase lack information about actual trends. A key motivation to improve CRVS systems is the concern about persistently high maternal and child mortality, but reliable information on mortality at all ages is essential to monitor progress toward achieving SDG 3 on health.

---

[15] Statistical Papers, Series M No. 19/Rev.3

[16] World Health Organization, *Improving Mortality Statistics through Civil Registration and Vital Statistics Systems: Strategies for Country and Partner Support*, 2014, http://www.who.int/healthinfo/civil_registration/CRVS_MortalityStats_Guidance_Nov2014.pdf?ua=1

[17] WHO, http://www.who.int/healthinfo/statistics/verbalautopsystandards/

[18] Alan Lopez and Philip Setel, "Better health intelligence: A new era for civil registration and vital statistics?, *BMC Medicine*, 2013, vol. 13, No. 73, pp. 1-4.

C. Household surveys

31. Household surveys, when based on representative probabilistic samples drawn from a frame that covers the whole population (such as a census), are the fourth key instrument for the generation of demographic data. Most household surveys are conducted face-to-face by trained interviewers. Household surveys collect detailed information that cannot be reliably obtained using other data-collection instruments.

32. In developing countries, demographic household surveys began to be used systematically in the 1970s to gather information about fertility and its determinants, including contraceptive use. In countries lacking reliable vital statistics, household surveys are the main source of data to estimate fertility and child mortality and, consequently, demographic surveys tend to be more common in those countries. Although surveys have also yielded data to estimate maternal and adult mortality at national level, or migration across geographic units, small sample sizes and reporting errors often produce unreliable estimates of these relatively rare events.

33. Household surveys are conducted more frequently than censuses but the frequency of those gathering demographic data varies considerably among countries. The sources of the most recent estimates of contraceptive prevalence show that 74 per cent of the countries in Africa carried out a survey during 2010-2014, compared to 57 per cent in Asia, 27 per cent in Latin America and the Caribbean, 10 per cent in developing Oceania and 19 per cent in developed countries.

34. International survey programmes have been major drivers of increased survey-taking. In the 1970s, the World Fertility Survey (WFS) programme set the standards for demographic surveys. Financed by UNFPA, USAID and the UK Overseas Development Administration, the program started in 1972 and had conducted 59 surveys when it ended in 1984, including 42 in developing countries. The WFS secretariat collaborated with national institutions to conduct and analyse the surveys in order to build national capacity. By producing model questionnaires with modules on specific topics and ensuring centralized data processing, the WFS achieved

harmonization of concepts and comparability of results across countries. Micro-data from WFS surveys for developing countries are still accessible in digital form.[19]

35. From 1977 to 1985, USAID funded the Contraceptive Prevalence Surveys (CPS) aimed at gathering information on contraceptive use by method and (public or private) source of services and supplies in order to evaluate and assist in managing family planning programmes. Forty-two CPS were conducted over that period.

36. The WFS and CPS set the stage for the next generation of international survey programmes. In 1984, USAID funded the first phase of the Demographic and Health Surveys (DHS), a programme still in operation that has provided essential data to monitor the MDGs.

37. In 1995, UNICEF launched a second major survey programme to monitor the implementation of the UN Convention on the Rights of the Child and the MDGs, the Multiple-Indicator Cluster Surveys (MICS). Although collecting mainly data relevant to monitor the health and well-being of women and children, the MICS also gather information on other topics. MICS belonging to the recently completed fifth round collected data to measure 130 indicators. By early 2015, 271 MICS had been conducted in 108 countries and micro-data from 152 were accessible online. UNICEF has developed a comprehensive set of tools, including questionnaire modules, to guide national survey teams through every step of the survey process, including planning, survey design, training of field teams, data collection, processing, documentation, analysis and dissemination.[20]

38. The DHS programme has also ensured comparability of data over time and across countries through standardization of modular questionnaires, data collection procedures and data processing. By the end of 2014, 252 standard DHS had been conducted in 89 countries or areas. Whereas DHS largely inherited the innovative approaches pioneered and tested by WFS in regard to demographic data collection, it has been innovating in the collection of health data. In 1995, anemia testing became a standard component of the DHS survey and since then, 20 other biomarkers have been added to DHS surveys in over 50 countries, including HIV and malaria testing. To address specific needs, DHS has developed specialized surveys on a range of topics, including surveys that provide an overview of a country's health service delivery.

[19] OPR Data Archive, http://opr.princeton.edu/archive/wfs/
[20] http://mics.unicef.org/tools

39. DHS was the first survey programme to use online interactive databases routinely to provide access to micro-data. Furthermore, DHS pioneered the systematic use of geo-referencing. In most of the recent DHS, groupings of households, denominated “clusters”, are geo-referenced using GPS, usually during the survey sample listing process. To maintain respondent confidentiality, GPS latitude/longitude positions are randomly displaced while ensuring that they remain within the area where each cluster is located. Therefore, DHS data can be integrated with other geo-referenced datasets, thereby leveraging their usefulness.

40. Both DHS and MICS have conducted some surveys via computer-assisted personal interviewing (CAPI) using tablets or laptops. The use of such devices can save time and enhance accuracy, but interviewers must be skilled in using the devices and sufficient technical support must be available to ensure quality control. Further innovation in this area is at the core of the survey project entitled Performance Monitoring and Accountability 2020 (PMA2020), whose objective is to monitor progress towards providing by 2020 family planning to an additional 120 million women who need it. The project is currently active in nine countries where over 1,200 female resident interviewers have been trained to conduct interviews at households and health-service delivery points, entering the data into Android smartphones equipped with a customized version of Open Data Kit (ODK). Interviewers transfer the data gathered to a central cloud-server via the mobile data network, thus enabling real-time data checking and fast data aggregation and analysis.

41. A catalogue of surveys gathering demographic information and conducted during 1995-2014 includes over 700 entries additional to those conducted by the major survey programmes discussed above. Three large sets are the 71 Living Standards Measurement Study (LSMS) surveys used to measure poverty by the World Bank, the 70 World Health Surveys conducted under the guidance of WHO in 2002-2004, and labour force surveys coordinated by the ILO.

42. Although most of the aforementioned surveys have been useful in palliating the lack of reliable vital statistics, they are no substitute for a well-functioning CRVS system. Demographic estimates derived from surveys are notoriously variable, owing both to random variation and reporting errors, particularly those involved in retrospective reporting of vital events. Yet, even if the availability and quality of vital statistics improve, demographic surveys will be needed to obtain

information on the determinants of demographic behaviour and on the array of indicators related to population necessary to monitor SDG implementation, such as those produced by DHS or MICS.

43. Longitudinal surveys are particularly useful to document changes over time and their causes. They can be multi-round surveys, where all persons residing in a sample area are covered, or panel surveys, where a particular group of individuals, rather than households, is followed from one round to the next. Panel surveys are being used to study population ageing. The Survey of Health, Ageing and Retirement in Europe (SHARE) covers 21 European countries and Israel. The Generations and Gender Programme (GGP) surveys 18-79 year-olds and has completed several rounds in ten countries and one round in another nine. Harmonized datasets are being made available . WHO's Study of Global Ageing and Adult Health (SAGE) covers persons aged 50 or over and some aged 18–49 in China, Ghana, India, Mexico, the Russian Federation and South Africa. In addition, 18 multi-round surveys and 13 panel surveys covering other topics were conducted during 1994-2014, including 16 in Africa and 10 in Asia.

44. Surveys will remain essential for SDG monitoring. In planning them, it must be remembered that individual surveys cannot cover many topics without compromising quality. If change occurs slowly, collecting data over short time intervals would be wasteful, and sampling variability may overwhelm the small changes one is trying to measure. Furthermore, survey samples are typically not large enough to represent small geographic areas or small population groups. The uneven geographic distribution of some sub-population groups relevant for the 2030 Agenda, including migrants and ethnic groups, may require special sampling techniques. In some cases, special modules may be added to existing household surveys. However, surveys alone cannot provide all the information necessary to ensure that nobody is left behind.

C. Health and Demographic Surveillance Systems

45. Health and demographic surveillance systems (HDSS) monitor demographic events and health in a population living in a well-defined geographic area. The process starts with a baseline census followed by periodic visits (one to four annually) to all households in the area to record events (births, deaths and

migration) and health-status changes. The system may also record marriages, divorces, changes in family status and household relationships. HDSS sites can be used to test interventions aimed at changing demographic behaviour or improving health, with some areas receiving the intervention and others used as controls. The repeated surveillance of sites improves the completeness of recorded events over time, since events missed in one visit will likely be caught at the next. Despite not being representative of the national population, HDSS sites have been generating useful demographic data, particularly on the age pattern of mortality and on causes of death via verbal autopsies in countries lacking a functional CRVS system.

46. There are two networks of HDSS. The ALPHA Network, coordinated by the London School of Hygiene and Tropical Medicine, consists of 10 sites in Eastern and Southern Africa devoted to the study of HIV/AIDS, including sero-surveillance.[21] The network maintains harmonized information from the 10 sites, but the data are not publically available.

47. The second network is INDEPTH,[22] established in 1998 and based in Accra, Ghana. It coordinates 52 HDSS sites in 20 countries in Africa, Asia and Oceania and is run by 45 research centres. INDEPTH coordinates multi-site projects generally focused on evaluating health interventions; acts as a clearinghouse for HDSS documentation and methodology, and maintains two public-access data repositories online.

III.Compilation, harmonization and dissemination

48. Some institutions and initiatives compile demographic data or catalogues to assist users in accessing the information available. One of them disseminates harmonized micro-data that are particularly valuable for comparative analysis.

49. Since 1948, the United Nations Statistics Division has been compiling and disseminating aggregate demographic statistics from censuses and CRVS via the yearly Demographic Yearbook. National data are obtained through a set of questionnaires dispatched annually to over 230 NSOs. Completed questionnaires can be submitted in digital form and full data series are accessible online.[23] UNSD maintains the World Census Programme website in which the implementation of

[21] alpha.lshtm.ac.uk
[22] www.indepth-network.org
[23] http://unstats.un.org/unsd/demographic/products/dyb/dyb2.htm

censuses is monitored closely. The website is complemented by the online Census Knowledge Database with technical publications, and a repository of census questionnaires worldwide. In Latin America, CELADE became the repository of census micro-data in the 1970s and, as noted above, facilitates the analysis of census micro-data by using REDATAM.

50. In 1999, the Partnership in Statistics for Development in the 21st Century (PARIS21) was founded to promote and facilitate statistical capacity development and a better use of statistics in developing countries. The increasing role that surveys had in monitoring the MDGs and the low utilization of survey results when they did not belong to international programmes led to the launch of the Accelerated Data Program (ADP) in 2006 to improve national survey programmes and increase the use of survey data.[24]

51. To that end, ADP provides technical and financial support for the documentation and dissemination of survey data. It supports the establishment of national survey databanks and the adoption of national data collection standards to foster comparability across data sources. ADP promotes open access to data and focuses on the documentation, preservation, dissemination, harmonization, collection and analysis of micro-data generated by censuses, surveys or administrative systems. ADP coordinates activities with the International Household Survey Network (IHSN), with which it shares a secretariat hosted by Paris21, and is financed primarily by the World Bank through a grant to PARIS21.

52. Established in 2008 by a grant from UK-DFID, IHSN provides technical and methodological guidelines for all stages of survey-taking, including standards, tools, and guidelines to document, disseminate, and preserve micro-data.[25] IHSN maintains a catalogue on the availability of survey and census data. The catalogue contains, when available, survey questionnaire(s), manuals and report(s), and a list of related citations. The catalogue uses NADA, a web-based cataloguing system that serves as a portal for users to search, compare or apply for access to census or survey information from low-income and lower-middle-income countries. The catalogue does not provide direct access to micro-data but, when available, provides a link to a site where the data may be accessed.

[24] http://adp.ihsn.org/
[25] www.ihsn.org

53. IHSN also maintains a list of other catalogues directing users to survey or census data. With respect to demographic data, the relevant entries in that list include the DHS site, UNICEF's Child Info for the MICS, the Global Health Data Exchange (GHDx) maintained by the Institute for Health Metrics and Evaluation (IHME) of the University of Washington, and IPUMS International[26] created by the Minnesota Population Center of the University of Minnesota. GHDx is a catalogue of surveys, censuses, vital statistics, and other health-related data.[27] It shows which data are available for download and which require a registration process or are available only upon request. It covers data sources for both developed and developing countries.

54. One notable source of demographic micro-data is IPUMS International. Begun in 1998 as a project to rescue census micro-data from destruction or deterioration, the Integrated Public Use Microdata Series International provides access to 277 census samples dating from 1960 to 2014 and covering 82 countries, three-quarters of which are developing countries. The census samples available are large: almost two-thirds include a least one tenth of the population and 85 per cent include at least one twentieth. For most countries, census samples are available for multiple years. The IPUMS datasets are harmonized to ensure comparability over time and across countries. Consistent naming conventions for variable are used and variables are recoded to facilitate comparability. Census documentation, including questionnaires, are provided in English and in their original language. IPUMS provides a user-friendly interface to obtain the required extracts from the data, thereby providing an invaluable service that has greatly increased the use of micro-data.

IV. Consistent and reliable estimates

55. The demographic evidence base encompasses not only data generated by statistical systems but also estimates derived from those data. Few institutions produce complete sets of consistent and reliable estimates of demographic indicators and, among them, the UN Population Division is a major producer. Ensuring consistency and reliability of the datasets produced requires an evaluation of the

[26] international.ipums.org/international/about.shtml
[27] http://www.healthdata.org/about/ghdx

basic data, adjustments for deficiencies detected, and the use of estimation methods appropriate for each situation.

56. The UN Population Division produces the most comprehensive set of population estimates and projections for all countries and areas of the world. These estimates and projections are revised biennially and are accessible online.[28] The population estimates are available by age and sex for every country and have been used by the whole United Nations system to derive MDG indicators requiring population inputs.

57. The Division also produces urban, rural and city population estimates and projections, urban and rural populations by age and sex, and estimates of the international migrant stock by age, sex, and countries of origin and destination. Estimates of adolescent fertility, contraceptive prevalence and unmet need for family planning have been produced by the Division in collaboration with UNFPA as MDG indicators. Also for the MDGs, the UN Inter-agency Group for Child Mortality Estimation, which includes the Population Division, has been producing estimates of infant and child mortality, and the Division collaborates with WHO, UNICEF, UNFPA and the World Bank in producing estimates of maternal mortality.

58. Two other initiatives have produced consistent sets of estimates of mortality and fertility referring mostly to developed countries. The Human Mortality Database (HMD),[29] created by the University of California at Berkeley and the Max Planck Institute for Demographic Research, contains time series of national life tables for 37 countries or areas having virtually complete vital statistics and census enumerations. It also contains data on registered deaths by age and sex, registered births and enumerated populations. A uniform method was used to construct the life tables, ensuring consistency over time and across countries. The HMD is accessible online and is the best source on changing mortality patterns in developed countries.

59. The Human Fertility Database (HFD)[30] is a project of the Max Planck Institute for Demographic Research and the Vienna Institute of Demography. It provides online access to cohort and period fertility estimates to monitor, analyse and forecast fertility in developed countries. National time series of fertility indicators, generally starting in the 1950s or before, are available for 26 countries, 24 of them developed.

[28] http://esa.un.org/unpd/wpp/
[29] www.mortality.org
[30] http://www.humanfertility.org/cgi-bin/main.php

The HFD is limited to populations with virtually complete birth registration and reliable population estimates over the reproductive ages. To ensure comparability, uniform methods are used to derive estimates.

60. The population estimates produced by the UN Population Division pertain to many of the special groups identified in the SDGs, including urban and rural populations. However, the estimation of urban and rural populations by age and sex is hindered by paucity of data since 28 per cent of countries lack such data, including 40 per cent of those in Africa. Similarly, 15 per cent of countries lack data on international migrants by age and sex. Although most countries collect some information on internal migration, variation in both the spatial and the temporal criteria used to identify internal migrants hinders the derivation of comparable estimates. ECLAC has developed three databases[31] of particular relevance for these topics: MIALC (Internal Migration in Latin America), IMILA (International Migration in Latin America) and DEPUALC (Spatial Population Distribution and Urbanization in Latin America). The project entitled Internal Migration around the Globe (IMAGE)[32] has made major advances globally in compiling data on internal migration (it holds data for about 100 countries or areas) and in developing indicators of internal migration that are comparable between countries.[33] The project has also considerably increased the availability on data and indicators on internal migrants by age and sex, estimates that are necessary to model population change at the sub-national level.

V. Integrating geo-referenced datasets

61. The integration of geo-referenced datasets, including those generated by satellites, to layer and map different types of information has been a major innovation. Such integration has proved essential in assessing the impacts of climate change, planning at the national and local levels, studying inequalities within countries and locating population groups vulnerable to natural disasters, famine and

[31] http://www.cepal.org/es/migracion

[32] https://www.gpem.uq.edu.au/qcpr-image

[33] Bell, Martin and Elin Charles-Edwards, Cross-national comparisons of internal migration: An update on global patterns and trends, UNDESA, Population Division, Technical Paper 2013/1.

other risks. Three institutions are the main producers of integrated geo-referenced datasets that include population indicators.

62. The Center for International Earth Science Information Network (CIESIN) of Columbia University was the first to overlay population data on a global raster.[34] The Gridded Population of the World, now in version 3 (GPWv3)[35], uses a raster with a 5 km^2 grid cell to redistribute the world population enumerated by censuses so that it is compatible with geo-referenced datasets of social, economic and environmental information.

63. The gridded dataset is constructed from national or sub-national units (usually administrative units) of varying resolutions. Separate grids are available for population counts and population density per cell, estimated for 1990, 1995 and 2000, and projected for 2005, 2010 and 2015. GPW comprises many maps, including of population density and sub-national administrative boundaries. The fourth version of GPW is under construction (GPWv4) and is based on the 2010 round of censuses. It will produce gridded datasets with population counts, population by age and sex, and population by urban/rural status using a 1 km^2 grid cell. Quinquennial estimates from 2000 to 2015 and a projection to 2020 are being prepared.

64. CIESIN also produces the Global Rural-Urban Mapping Project (GRUMPv1), which has three components: a gridded population data at higher resolution than GPW for 1990, 1995, and 2000; a dataset delineating urban areas on the basis of night-lights data produced by the National Oceanic and Atmospheric Administration or from buffered settlement centroids where night lights are not sufficiently bright, and a dataset with a point for each urban area with at least 1,000 inhabitants. GPW and GRUMP have been used to estimate the populations at risk of specific natural disasters.

65. Using geo-referenced data on infant mortality and proportion of underweight children from DHS and MICS, CIESIN has produced high-resolution maps to detect “hotspots” of poor child health. Layering environmental maps over the GPW, CIESIN has generated maps showing the population densities of areas highly vulnerable to climate change.

[34] A raster consists of a matrix of cells (or pixels) organized into rows and columns (or a grid) where each cell contains a value representing information, such as population counts.

[35] http://sedac.ciesin.columbia.edu/data/collection/gpw-v3

66. The WorldPop project of the Geo-Data Institute of the University of Southampton generates open access datasets on the spatial distribution of populations in Latin America, Africa and Asia to support development planning, disaster response and health applications. It produces updatable maps with accompanying metadata and measures of uncertainty. More recently available remotely-sensed and geo-referenced datasets, such as settlement locations, settlement extents, land cover, roads, building maps, health facility locations, satellite night lights, vegetation, topography and location of refugee camps, have been used to model population densities at a high spatial resolution. The resulting gridded population datasets have been the basis for deriving spatial distributions of other population indicators, such as births, pregnancies and population by age.

67. IPUMS has been developing TerraPop, a system that integrates three types of geo-referenced data: population micro-data from IPUMS; summary indicators for geographic areas, including economic indicators or policy information; and global raster data derived from satellite imagery and climate models. The first phase of TerraPop involved geo-referencing existing census micro-data to maintain the consistency of administrative boundaries over time, thus enhancing the inter-temporal comparability of sub-national data. TerraPop will allow users to convert micro-data to raster format for visualization or spatial modelling.

VII. “Big Data” as a complement to demographic data

68. Because of the widespread penetration of ICT, digital information is being generated continuously by GPS devices, mobile phones, automated teller machines, scanning devices, sensors, satellites, search engines and social media. The large volume, high velocity and wide variety of the data thus generated are defining characteristics of what is called “big data”. Most big data are available in real-time and with minimal cost.

69. However, big data are often unstructured and do not fit well into relational databases. Being massive, their analysis demands considerable computer power, deep knowledge of ICT and of the data-generation process. Lack of representativity is a major drawback of big data since any insights they yield cannot be attributed to the general population or even to a well-defined population group. Furthermore,

most big data are in private hands and legislation designed to permit their use for statistical purposes has not yet been adopted in most countries. Issues of concern about privacy and confidentiality need to be addressed if big data are to become a source for regular statistical analysis.

70. So far, satellite-imaging data have been the main type of big data used for demographic applications and, as documented above, static datasets have been used rather than the continuous flow of satellite imaging. Call detail records (CDR) generated by mobile phones have been used to explore human mobility and relate it to different processes, including natural hazards, road traffic and the spread of infectious diseases. CDR have also been useful in mapping populations at risk of natural hazards, climate change or epidemics. Applications that relate population indicators to CDR in order to estimate the population in raster cells or other small areas show some promise. However, use of CDR at a global scale is not yet a reality, partly because no single company can provide CDR with complete global coverage.

71. WorldPop is planning to use the geo-location of social media activity as an additional variable in the model used to redistribute population in a raster and to assess whether these data can improve projections of geo-located populations. So far, all studies relating population indicators to big data use geo-referenced demographic information from censuses or surveys as necessary inputs. CDR and tweets are at best a correlate of population indicators and therefore far from being able to substitute for traditional demographic data sources.

VIII. The challenge of disaggregation

72. Monitoring the SDGs will require population data disaggregated by age and sex, race or ethnicity, migratory status, disability, income and geographic location. The three demographic data systems that cover the whole population—censuses, population registers and CRVS—can produce the required level of disaggregation provided they gather the relevant information, geo-reference the data at the smallest administrative level or higher spatial resolutions, and provide access to micro-data for the full population so that information for particular groups can be extracted. If access is limited to samples, data producers should release aggregate population numbers based on full counts and classified by relevant socio-economic

characteristics, including sex and single years of age, for the smallest administrative areas.

73. Countries having population registers can link them to other administrative datasets. However, because many countries lack population registers or complete CRVS—the only systems that produce continuous coverage of the whole population—modelling will be needed to estimate disaggregated populations over intercensal periods. Surveys will continue to fill gaps in demographic information caused by incomplete vital statistics and to measure indicators that other sources do not yield. Geo-referencing survey data routinely to integrate such information with other datasets will boost their utility and permit geographical disaggregation by using statistical modelling.

IX. The way forward

74. The availability of reliable and timely demographic data is essential for planning and implementing interventions to achieve the SDGs and to monitor progress toward their achievement. As this report has shown, much remains to be done to strengthen data systems and to support the in-depth analysis, disaggregation, and use of the data generated. The following recommendations provide guidance on how to achieve those ends.

75. Donors should support the Global Financing Facility for Reproductive, Maternal, Newborn, Child, and Adolescent Health (GFF) or other financing mechanisms that provide assistance for the establishment or expansion of **CRVS** systems in developing countries.

76. Governments of countries lacking complete vital statistics should commit to establish a functioning CRVS system and ensure that it generates timely vital statistics. Governments of countries with incomplete or deficient vital statistics should develop and implement strategies to improve coverage and reporting accuracy, and to foster the collaboration of civil registration with the health system to improve the reporting of causes of death according to the guidelines set by WHO.

77. Governments should support the preparation and undertaking of the **2020 round of censuses**. National authorities should leverage technology to improve the efficiency of data collection, processing and dissemination. They should ensure that

geo-referencing becomes standard practice in census-taking and that the corresponding spatial datasets are made available in digital form.

78. Governments should adopt **open-data policies** allowing the dissemination of public use, geo-referenced and anonymized micro-data from censuses, demographic surveys and CRVS.

79. National authorities should ensure that the standards, guidelines and tools used for data collection are consistent with guidance provided by the UN Statistical Commission and other relevant international or regional organisations and networks.

80. Data producers should ensure that **demographic statistics are classified by age** and are presented by single year of age from 0 to 99 and always by sex.

81. Governments and donors should support and promote **NSO's access to the necessary ICT** equipment and software to improve the collection, processing and dissemination of data. Equally important is to support personnel-training in the use of new technologies.

82. Governments should **strengthen the capacity of NSOs to conduct household surveys** and encourage data producers to follow the guidelines developed by ADP and IHSN on how to conduct household surveys, document the survey process, disseminate the data gathered and archive them. Specialized surveys may be needed for population subgroups that are unevenly distributed in the national territory, including internal and international migrants.

83. Governments should encourage and **support efforts to integrate demographic and other data sources via geo-referencing**, including by facilitating access to satellite images. This is particularly important given the emphasis of the 2030 Agenda on the well-being of all population groups, reinforcing the need for data disaggregation and for moving beyond national averages. The private sector should support such efforts by providing access to big data under its control.

84. Governments and donors should continue supporting work on **harmonization and data-integration** and are encouraged to contribute micro-data from the 2010 round of censuses to national, regional and international repositories in order to facilitate their use in setting a baseline for SDG monitoring.

85. **Governments should develop policies for the use of big data**, including safeguards to maintain privacy and confidentiality, and guidelines to facilitate access to data controlled by private firms.

86. Governments should encourage **greater cooperation between national research institutions and NSOs** to foster the in-depth analysis of existing demographic data. Partnerships with institutions already involved in producing globally integrated datasets or consistent and comparable sets of demographic estimates should also be fostered.

87. The **United Nations system**, including the Secretariat, the regional commissions and specialized funds, should continue their work and support to the inter-agency and expert groups to harmonize estimates, improve estimation methods, report progress at the regional and global level, and enhance national capacities in the production and use of quality demographic data.