



Swiss National Cohort (SNC) and Longitudinal demographic statistics (DVS)

Integration of the Swiss National Cohort (SNC) into the Longitudinal Demographic Statistics (DVS)

Neuchâtel, January 2023

Published by: Federal Statistical Office (FSO)
Information: info.pop@bfs.admin.ch, tel. +41 058 467 25 25
Editors: Nathalie Monnet, FSO
Contents: Nathalie Monnet, FSO
Topic: 01 Population
Original text: French
Translation: FSO language services

Layout concept: Section DIAM
Downloads: www.statistics.ch
Copyright: FSO, Neuchâtel 2023
Reproduction with mention of source authorised
(except for commercial purposes)

Table of contents

1	Current status	3
2	SNC-DVS integration	3
3	Reservations regarding the SNC-DVS integration	4
4	Conclusion	5

1 Current status

The Swiss National Cohort (SNC) is a database that was produced by the Epidemiology, Biostatistics and Prevention Institute of the University of Zurich and the Institute of Social and Preventive Medicine of the University of Bern (ISPM - the project's implementing body). The SNC project links individuals in the 2010-2011 population and households statistics with data from the 1990 and 2000 federal population censuses through *probabilistic matching*¹.

Since January 2022, the FSO has been in charge of the SNC data. The database includes:

- The SNC extract which contains all individuals present in the 1990 census (FPC 1990), the 2000 census (FPC 2000) and STATPOP (population and households statistics) from 2010 to 2019. Some of the variables come from several other statistics: the vital statistics (BEVNAT) for the date of death and causes of death, the structural survey (SE) and the buildings and dwellings statistics (BDS), the 1992 Swiss health survey (SHS) and a socio-economic index created by the ISPM.
- A key file that allows the use of the probabilistic link between the 13-digit OASI number of individuals in STATPOP and the censuses' identifiers.

The Longitudinal Demographic Statistics (DVS) contain all movements from 2011 onwards of persons present in at least one STATPOP file as well as an attribute table with variables that do not change over time (date of birth, place of birth, parent/child relationship). In addition, population records on December 31st of each year have been created on the basis of the attribute and movements tables to assist users in their analysis.

On the one hand, the SNC contains population records in 1990, 2000, and from 2010 to 2019. On the other, the DVS cover the part from 2010 onwards, except for the causes of death and the structural survey variables². Integrating information from the 1990-2000 censuses into the longitudinal demographic statistics would make it possible to extend the biographies of individuals (for example, by adding information on their commune of residence in 1990 and 2000) as well as adding individuals who are no longer present from 2010 onwards. The necessary steps are detailed in the second section. However, there are several limitations to this integration that need to be taken into account, which are detailed in the third section.

2 SNC-DVS integration

For the integration of the 1990 and 2000 censuses' information into the DVS, the following steps are necessary:

1. Ensure the quality of the SNC probabilistic matching, e.g., by (1) updating the OASI number in the 1990 and 2000 data; (2) looking at how many people returned to Switzerland after 2012 (if absent in 2010-2011) who could have been present in 1990 and 2000, as the link was not made in the SNC project. Other techniques could also be used to test the links.
2. Resume the date of birth and place of birth in the attribute table (DVS) for persons who are in the 1990 and 2000 censuses.
3. Harmonise the definition of the place of residence: the censuses are based on the population at the economic residency, while the DVS is based on the population at the main place of residency.
4. Create movements between 1990 and 2010: only births, deaths in Switzerland and acquisitions of Swiss citizenship could be created but without any information on the other variables at the time of the event (commune of residence, nationality, marital status, etc.). Migration, relocation, change of permissions to stay, change of population type, change of marital status, etc. would not be feasible.

¹ The censuses do not have OASI numbers but a unique identifier for each individual that differs for each census.

² This information can be obtained by requesting a data linkage between the DVS and the various files that enabled the creation of the SNC project (causes of

death, the building and dwellings statistics, the structural survey, etc.): [Data linkages for third parties | Federal Statistical Office \(admin.ch\)](#).

3 Reservations regarding the SNC-DVS integration

Despite the biographical contribution for the 1990s and 2000s, there are several arguments against incorporating SNC data into the longitudinal demographic statistics:

1. Few requests: In 2022, only four requests solicited the SNC + FPC + STATPOP data (data linkages requests). Few requests require data from before 2010 and we believe that what is currently available is sufficient. Indeed, the SNC data are of good quality by themselves, i.e., they are sufficient for the requests received.
2. Conceptual differences: Censuses include different definitions of the type of population (economic population). The longitudinal demographic statistics are based on the population at the main residence. Moreover, the SNC only includes two states (1990 and 2000) whereas the DVS is an event statistic (all movements from 2011 onwards). A great deal of consistency would be needed, without adding much additional information. Indeed, the years 1991-1999 and 2001-2009 are not available (e.g., links with BEVNAT have not been established).
3. Quality uncertainties: The probabilistic matching carried out by the ISPM to find the OASI number of individuals present in the 1990 and 2000 censuses is based on individuals present in STATPOP between 2010-2011. A potential problem arises from individuals present in one of the censuses and in STATPOP except between 2010-2011. For example, if a person was in Switzerland in 2000 and 2014 but not in 2010-2011, no link has been established for this person and they appear as two separate persons in the SNC.

Other quality problems are related to the difficulty of matching individuals with a 10-year gap, especially for people aged 15-25 in 1990 or 2000. Indeed, the probability of change of marital status, commune of residence, etc. is higher for this category than for other age groups. The data show a spike in missing links between censuses' identifiers and the OASI number for people born in the 1980s. These quality problems do not allow for an optimal match with the work done for the longitudinal demographic statistics, which is based on deterministic links using the OASI number.

4 Conclusion

In conclusion, the arguments put forward are against the integration of the SNC into the DVS. At present, the DVS does not yet cover all the themes covered by the STATPOP data. The work of creating a comprehensive product is initially the priority for longitudinal demographic statistics. In addition, all the information from the integration of the SNC into DVS is already available in its current state, as shown in the diagram below, using the key file. Not integrating the SNC into the DVS does not prevent the two files from being linked. This is because the two files are not based on the same types of link-age (SNC based on probabilistic matching, DVS based on deterministic linkage based on the OASI number).

According to the above reservations, it is not advisable to take resources for work that does not seem to add substantial value to the current situation.

