



Securing High-quality Data on Populations

Why we need EuroCohort, GGP and SHARE in Europe

Key messages:

- ▶ Research infrastructures such as EuroCohort, GGP and SHARE are key to ensure high quality, evidence-based social policies in Europe.
- ▶ EuroCohort, GGP and SHARE can deliver a comprehensive picture of the European life course from early years until later life. They cover all the major phases of life and adapt the methods and content of their data collection to suit the target population.
- ▶ These research infrastructures closely collaborate in order to ensure synergies, comparability in the data collected and deliver efficiencies and innovations in data collection processes.

Authors:

Tom Emery
Gary Pollock
Annette Scherpenzeel

Editor:

Daniela Vono de Vilhena

Introduction

Evidence-based policy requires high-quality data. Fortunately, we are living through a data revolution, which is opening up new opportunities for better quality data to feed into the policymaking process. There is an increasingly diverse range of data to help inform policy. Data from censuses and population registers, as well as from surveys, biomarkers, digital trace data and genetic data can help us triangulate and deepen our understanding of populations. However, when there is such a vast amount of data available, there is a danger that we end up drowning in numbers. To prevent this, research infrastructures have been created to collect, process, organise and disseminate high-quality data. These research infrastructures are governed by what scientists call the FAIR principles:

- **Findable:** The data need to be findable by researchers and policymakers. Data from research infrastructures need to make sure that the collected data is documented and catalogued in archives and that users can find the specific information they need.
- **Accessible:** The research infrastructures need to provide ways for users to access the data. For some data, this is quite straightforward as they only need to verify that the user is a researcher. However, some data is sensitive, so research infrastructures need to provide ways for users to securely access this data.
- **Interoperable:** The research infrastructures are responsible for making sure that data is interoperable and comparable with different sources. They do this by using industry standards and classifications so that data can be linked to many different sources.
- **Reusable:** Once the data has been analysed by users, the research infrastructures need to make sure that the analysis can be reproduced by other researchers. This helps improve accountability in data analysis and facilitate learning amongst the user community.

There are three research infrastructures which have been created to implement the FAIR principles in the study of European populations and ensure that policymakers and researchers have access to the best data possible. Each study focuses on a separate part of the life course. The **EuroCohort** study focuses on generating data on children and adolescents. The **Generations and Gender Programme (GGP)** collects data on young adults and midlife. The **Survey of Health, Ageing and Retirement in Europe (SHARE)** covers people in later life, up until death. Whilst they each centre around a separate area of the life course, the infrastructures collaborate to help improve the quality and quantity of data available to researchers

through new data innovations and collaborations.

EuroCohort

The EuroCohort study is the youngest of the three infrastructures and is not yet operational. It intends to implement a child cohort study in all European Member States and follow these children over 25 years to see how they develop. It is broadly based on the British Cohort studies such as the 1970's British Cohort Study and the Millennium Cohort Study. The study is currently coordinated by the Manchester Metropolitan University and is just completing its design study phase, in which the questionnaire and sampling strategy are provisionally agreed upon.

The aim of the EuroCohort study will be to track multiple dimensions of children's wellbeing, as well as key indicators of their background and immediate context. Health and educational outcomes, cognitive ability and personal wellbeing will be measured along with detailed information on housing, family, educational and social surroundings. The questionnaire will develop and adapt as the children grow up to reflect their changing lives, particularly as they move through the education system and develop into young adults. The Cohort study is key to better understand how family and social contexts shape young people's outcomes and therefore better inform policies in these areas. Its longitudinal design will enable researchers to better identify the causes of inequalities and how they evolve over time, while its cross-national and comparative design will allow for monitoring of European trends over time, and to identify what policies do and do not work.

The impact of existing cohort studies has illustrated the potential for a pan-European Cohort study. They have had considerable impact in shaping educational, health and family policies, particularly in the earliest years of life. Research has persistently shown that early life experiences have a clear effect on later life outcomes and the EuroCohort study will be able to shed further light on these processes.

GGP

The Generations and Gender Programme (GGP) was established in 2001 and collects data on relationships, fertility and family dynamics in adulthood, as well as individuals' plans for the future. These are all vital for understanding population trends such as low fertility and complex family forms, but also social issues like social mobility and in-

equality. The project was initially coordinated by the United Nations Economic Commission for Europe (UNECE) and has a broad geographic scope. Since 2009, the Netherlands Interdisciplinary Demographic Institute (NIDI) in the Hague has coordinated the GGP.

In the new round of data collection planned for 2020, the GGP questionnaire will include a new range of questions that will help to better identify the changing nature of work including flexibility, work-life balance, job security and questions on mobility and migration. It will also include questions on the digitisation of the life course and the extent to which online dating and the use of digital contact methods are reshaping family life.

The GGP recently developed the option to complete the survey online, reducing the costs of fieldwork substantially and opening up new possibilities for gathering data. In the new round of data collection, for the first time, many countries will field the survey primarily online. This will enable rapid collection and processing of data from a wide range of countries.

The GGP has had a wide ranging impact on social and population policies in Europe. It has over 4,000 users worldwide who have produced over 1,400 publications. The GGP works particularly close with statistical offices and demographic institutes in many countries to ensure that the microdata collected is fed into the policymaking process on school-to-work transitions, work-life balance and family leave policies.

SHARE

The Survey of Health, Ageing and Retirement in Europe (SHARE) is a landmark of the European Strategic Forum for Research Infrastructures and was the first European Research Infrastructure Consortium. SHARE has collected data over nearly two decades, following individuals in 28 countries through the ageing process. The data collected by SHARE include detailed and high-quality information on health and finances of individuals in later life and helps policymakers understand the ageing process at a time of accelerating demographic change.

With a rapidly ageing population, understanding the ageing process is more vital than ever and essential for developing sustainable social policies. SHARE collects data on the health, social and economic situations of people in later life. This includes a detailed medical history, as well as measures of their current health like biomarkers, grip strength

and cognitive ability. Social networks and social contacts of individuals are also captured to provide a wider social context to individuals' lives. Finally, it includes a detailed breakdown of individuals' financial status including their pension, their assets and the costs associated with health and care. By following trends over time, it is possible to see how an individual's wellbeing develops during later life and to identify determinants of poor health and financial vulnerability.

SHARE data is collected via face-to-face interviews which last approximately one hour. These interviews take significant resources, but are necessary to collect the type of data required for the study of ageing. For example, health indicators are vital for the study of later life and it is therefore necessary that SHARE interviewers are highly trained and spend significant time with respondents in order to collect the relevant health information. Interviewers also help respondents provide detailed information about personal finances, labour market activity and other economic measures, which are often hard for respondents to answer independently.

The impact of SHARE data has been vast. Over 10,000 users worldwide use the data and over 2,500 publications have been produced to date on a wide range of topics. The cross-national and longitudinal nature of the data have been particularly useful in illustrating the way in which social context shapes outcomes in later life and have been used to evaluate and design a wide range of public policies including pension and public health policy.

Making Life Course Data FAIR

The three research infrastructures closely collaborate on a wide range of initiatives to make population data FAIR. The GGP has utilised a range of tools developed by SHARE to better standardise and document the data it collects for the use of researchers. For example, the GGP has adopted a tool that helps to better classify an individual's occupation and make it comparable with the data collected by SHARE. Similarly, the GGP and SHARE work closely together to ensure that they have access to the highest quality sampling frames on which to conduct their surveys. These sampling frames are the lists of people which surveys use to select respondents, and data owners are often reluctant to provide access to researchers. By working together, data access and data quality was significantly improved in both studies.

The GGP has been working closely with the EuroCohort

study as part of its design work to ensure that they implement the highest methodological and management standards. For example, the GGP has experience in online data collection, which will be an important component of EuroCohort in later waves. The GGP shares technical specifications and advises the EuroCohort study on design components. By doing so, it ensures that EuroCohort will not have to reinvent the wheel for many technical challenges. Instead, EuroCohort can concentrate investment on collecting the highest quality data possible.

The EuroCohort, GGP and SHARE also work together to align questions and measures across the three infrastructures so that researchers can observe various dynamics across all stages of the life course. For example, as mentioned earlier, both SHARE and the GGP have shown that early life circumstances play a role in shaping outcomes in later life. Given this, the three infrastructures work to ensure that the same definitions and measures of early life circumstance, such as parental divorce, material deprivation or poor health, are used consistently.

Despite this collaboration however, the three infrastructures face many shared challenges. First, it is very difficult for the surveys to make optimal use of administrative data from multiple countries. There are many technical, legal and ethical challenges to doing this which the GGP and SHARE have explored as part of SERISS (Synergies for Europe's Research Infrastructures in the Social Sciences, www.seriss.eu), a joint project on developing social science research infrastructures. Yet, significant work remains before effective solutions that link to administrative data and government records is in place. Second, all three infrastructures rely on longitudinal and cross-national data collection. Failure to collect data at a given time point or in a given country, can seriously weaken scientific power. Maintaining funding in 28 countries over decades is exceptionally difficult and poses the single biggest risk to all three studies.

Policy Recommendations

- Serious investment is required in the collection of data on populations for research purposes. These data help develop an understanding and better policies regarding the key demographic challenges facing Europe today. Without such investments, key policy decisions will be made without all the facts.

- Research infrastructures provide the data we need to develop better policies. To support the development of infrastructures for the study of populations, the GGP and

EuroCohort project should be admitted to the European Strategic Forum for Research Infrastructures.

- To support better policymaking, statistical offices and governments should work to make administrative data available for linkage with research infrastructures. SERISS has produced recommendations on how this can be done for both survey sampling and to enrich survey data.

- A new European funding mechanism is required to ensure the stability and continuity of longitudinal data collections in European Member States. This funding mechanism could be through Horizon Europe and would ensure that the research infrastructures that study populations are truly European in coverage. ■

Imprint

Publisher: Max Planck Society for the Advancement of Science on behalf of the collaborative network "Population Europe"

Technical Coordination: Emily Lines

Layout: The Brettingshams GmbH, Berlin; Karen Olze / Judith Miller / Produkt und Grafik Design

Photo Credit (Cover Page): ©Natee Meepian – iStock

Print: Newprint blue GmbH, Berliner Str. 13-14, 10715 Berlin

ISSN: 2512-6164

Contact: Population Europe Secretariat, Markgrafenstraße 37, 10117 Berlin, Germany

Phone: +49 (0)30 2061 383 30, **Fax:** +49 (0)30 2061 383 50

Email: office@population-europe.eu

Web: www.population-europe.eu

The opinions of the authors do not necessarily reflect those held by the publisher or the editorial office. Reprints of any part of the articles are permitted as long as the original authors and publishers are given credit. Please provide us with a specimen copy.



This publication is funded by the Horizon 2020 Research and Innovation Programme under grant agreement n°739511 for the project "Generations and Gender Programme: Evaluate, Plan, Initiate".