Program for Studying Intergenerational Transmissions in Infant Mortality Using the Intermediate Data Structure (IDS)

By Luciana Quaranta

To cite this article: Quaranta, L. (2018). Program for Studying Intergenerational Transmissions in Infant Mortality Using the Intermediate Data Structure (IDS). *Historical Life Course Studies, 7*, 11-27, http://hdl.handle.net/10622/23526343-2018-0010?locatt=view:master

HISTORICAL LIFE COURSE STUDIES

Intergenerational transmissions of infant mortality using the Intermediate Data Structure (IDS)

VOLUME 7, SPECIAL ISSUE 2
2018

GUEST EDITORS
Luciana Quaranta
Hilde L. Sommerseth

EHPS NETWORK

MISSION STATEMENT

HISTORICAL LIFE COURSE STUDIES

Historical Life Course Studies is the electronic journal of the *European Historical Population Samples Network* (EHPS-Net). The journal is the primary publishing outlet for research involved in the conversion of existing European and non-European large historical demographic databases into a common format, the Intermediate Data Structure, and for studies based on these databases. The journal publishes both methodological and substantive research articles.

Methodological Articles
This section includes methodological articles that describe all forms of data handling involving large historical databases, including extensive descriptions of new or existing databases, syntax, algorithms and extraction programs. Authors are encouraged to share their syntaxes, applications and other forms of software presented in their article, if pertinent, on the EHPS-Net website.

Research articles
This section includes substantive articles reporting the results of comparative longitudinal studies that are demographic and historical in nature, and that are based on micro-data from large historical databases.

Historical Life Course Studies is a no-fee double-blind, peer-reviewed open-access journal supported by the European Science Foundation (ESF, http://www.esf.org), the Scientific Research Network of Historical Demography (FWO Flanders, http://www.historicaldemography.be) and the International Institute of Social History Amsterdam (IISH, http://socialhistory.org/). Manuscripts are reviewed by the editors, members of the editorial and scientific boards, and by external reviewers. All journal content is freely available on the internet at http://www.ehps-net.eu/journal.

The European Science Foundation (ESF) provides a platform for its Member Organisations to advance science and explore new directions for research at the European level. Established in 1974 as an independent non-governmental organisation, the ESF currently serves 78 Member Organisations across 30 countries. EHPS-Net is an ESF Research Networking Programme.

The European Historical Population Samples Network (EHPS-net) brings together scholars to create a common format for databases containing non-aggregated information on persons, families and households. The aim is to form an integrated and joint interface between many European and non-European databases to stimulate comparative research on the micro-level.
Visit: http://www.ehps-net.eu.

HISTORICAL LIFE COURSE STUDIES
VOLUME 7, SPECIAL ISSUE 2 (2018), 11-27, published 24-05-2018

Program for Studying Intergenerational Transmissions in Infant Mortality Using the Intermediate Data Structure (IDS)

Luciana Quaranta
Lund University

ABSTRACT

Studies conducted in historical populations and developing countries have evidenced the existence of clustering in infant deaths, which could be related to genetic inheritance and/or to social and cultural factors such as education, socioeconomic status or parental care. A transmission of death clustering has also been found across generations. One way of expanding the knowledge on intergenerational transfers in infant mortality is by conducting comparable studies across different populations. The Intermediate Data Structure (IDS) was developed as a strategy aimed at simplifying the collecting, storing and sharing of historical demographic data. The current work presents two programs that were developed in STATA to construct a dataset for analysis and run statistical models to study intergenerational transfers in infant mortality using databases that are stored in the IDS. The programs use information stored in the IDS tables and after elaborating such information produce Excel files with results. They can be used with any longitudinal database constructed from church books, civil registers, or population registers.

Keywords: Historical demography, Intergenerational transmissions, Infant mortality, Intermediate Data Structure, STATA, Comparative research

e-ISSN: 2352-6343
PID article: http://hdl.handle.net/10622/23526343-2018-0010?locatt=view:master
PID software: http://hdl.handle.net/10622/23526343-2018-0011?locatt=view:master
The article and the software can be downloaded from here.

1 INTRODUCTION

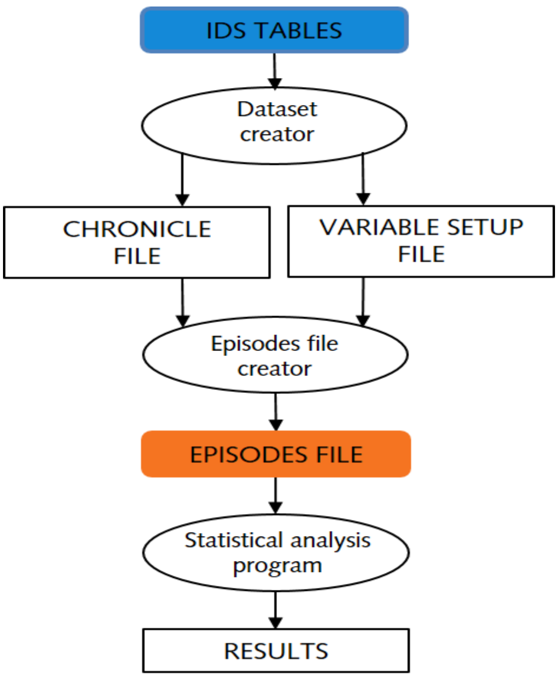
Previous research has shown the existence of infant mortality clustering within certain families (Das Gupta, 1990; Edvinsson, Brändström, Rogers, & Broström, 2005; Janssens, Messelink, & Need, 2010; Vandezande, 2012), possibly resulting from genetic inheritance and social and cultural factors related to education, socioeconomic status or parental care (Janssens et al., 2010). A transmission of death clustering has also been found across generations (Vandezande, 2012). One way of expanding the knowledge on intergenerational transfers in infant mortality is by conducting comparable studies across different populations.

The use of longitudinal micro level historical demographic data presents many challenges, which are often connected to their multilevel and relational aspects as well as to the complexity of conceptualizing processes that develop over time. The Intermediate Data Structure (IDS) was developed as a strategy aimed at simplifying the collecting, storing and sharing of historical demographic data (Alter & Mandemakers, 2014; Alter, Mandemakers, & Gutmann, 2009). The structure provides a common platform to store data from different databases, regardless of their original form. Among other advantages, the structure facilitates the sharing of data and software and it increases the transparency of how data outputs are prepared.

Taking advantage of the possibility to conduct comparative research provided by the IDS, a project aimed at studying intergenerational transmissions in infant mortality across five different populations – in Belgium (Donrovich, Puschmann, & Matthijs, 2018), the Netherlands (Van Dijk & Mandemakers, 2018), Norway (Sommerseth, 2018), Northern Sweden (Broström, Edvinsson, & Engberg, 2018) and Southern Sweden (Quaranta, 2018) – was started. The current article presents two programs developed within this project to create the dataset for analysis and run statistical models. These programs have been used in the study of each of these five populations, as well as in an analysis using a pooled dataset of all such areas (Quaranta et al., 2017). Besides the interest per se on expanding our knowledge about the determinants of infant mortality, the current work has also the scope of evidencing the great advantages of adopting the IDS for research using longitudinal historical demographic databases.

The programs presented in this work are generically written and they can be used with any longitudinal database that has been transposed into the IDS. In fact, they work for databases created from church books, civil registers, or population registers. Databases constructed from church books or civil registers generally consist of family reconstitutions made from records of births (or baptisms), deaths (or burials) and marriages. Population registers, in addition to vital events, contain continuous information about migration and the family structure. Whereas, in datasets originating from population registers, it is possible to know exactly when an individual entered or exited from the study sample. In databases created from church books or civil registers, data are of a snapshot character and only known from specific observed events.

Figure 1 *Workflow of the programs created to study intergenerational transmissions in infant mortality with IDS data*



The workflow of the programs is presented in Figure 1. The first program developed specifically for this project, the "Dataset creator", elaborates basic information stored in the IDS tables to produce variables and create the dataset for analysis. The "Episodes file creator", a generic IDS software which has been previously developed (Quaranta, 2016), transforms the data extraction into a rectangular file for analysis. The second program developed within this project, the "Statistical analysis program", is devoted to the empirical analysis and estimates descriptive statistics and statistical models and saves such results in Excel files.

Through the use of the "Dataset creator" and "Statistical analysis program"[1] it is possible to easily and automatically produce graphs and tables of results that can be directly inserted into a research article. The programs were written in STATA but the dataset constructed for analysis can also be imported into other statistical software in order to conduct further analysis using those packages. By using the command 'saveold' in STATA it is also possible to save the data file to older versions of the program.

The current paper is structured into different sections. The second section discusses how data can be stored into the IDS and extracted from such structure. The third section presents in more detail the "Dataset creator". The fourth section discusses in detail the "Statistical analysis program". The last section contains discussions and conclusions.

2 THE STORAGE AND EXTRACTION OF DATA USING IDS

The IDS consists of five main tables: INDIVIDUAL, which is used to store information relating to individuals; INDIV_INDIV, which defines relationships between individuals; CONTEXT which contains information about physical or social places or environments that affect one or more persons, defining the characteristics that are shared by groups of individuals; CONTEXT_CONTEXT which defines the relations between different layers in a hierarchy of contexts; INDIV_CONTEXT, which contains information on spells of time during which individuals are present in a specific context (Alter & Mandemakers, 2014). The main fields included in the INDIVIDUAL table are Id_I (an identifier of the individual), Type (an identifier of the Type of attribute) and Value (the value of the variable), and the time stamp (Day, Month, Year). The IDS also contains a METADATA table, which is used to describe the variables and values included in the five main tables. Data stored in IDS follows the Entity Attribute Data Model (Stead, Hammond, & Straube, 1982), and therefore contains one attribute per each record. In other words, each row of the table has one declaration of a Value of a Type. A detailed description of how to store data into IDS is given in Alter and Mandemakers (2014).

The IDS is intended for the storage of data obtained from sources (e.g., date of birth, individual occupation, date of marriage, address) or from existing databases that had already organised these kind of data in other formats. Using the Extended Intermediate Data Structure (EIDS) variables that are constructed from source information that is stored in the IDS (e.g. household size, head occupation, presence of parents) can also be stored, therefore allowing researchers to use such variables in different studies (Quaranta, 2015). The EIDS has the same structure as the IDS and is based on the tables INDIVIDUAL_EXT and CONTEXT_EXT. The METADATA table of IDS can be expanded to also include information on the definition of variable Types and Values of data stored in the EIDS.

In order to conduct statistical analysis using longitudinal demographic data, the dataset needs to be formatted as an EPISODES FILE, which is a rectangular data array that is ready for statistical analysis. An EPISODES FILE contains one or more spells per each individual. Spells are periods of time during which the values of variables remain constant and at the end of the event of interest the study can take place. Each row of the EPISODES FILE has a start and an end date.

Data extractions made from the IDS and EIDS table should be included in the so-called CHRONICLE FILE and VARIABLE SETUP FILE, which are used to create the EPISODES FILE (Quaranta, 2015). The CHRONICLE FILE contains all the variables selected for analysis, which can be a combination of individual or context level time-varying variables, time-invariant variables and events stored in the IDS and the EIDS tables or created by extraction programs from IDS data and added directly. The CHRONICLE FILE also follows the Entity Attribute Data Model, and contains the fields Id_I, Type, Value, Day, Month, Year, and DayFrac. The field DayFrac is aimed at handling date collisions, which occur when there is more than one Value of a specific

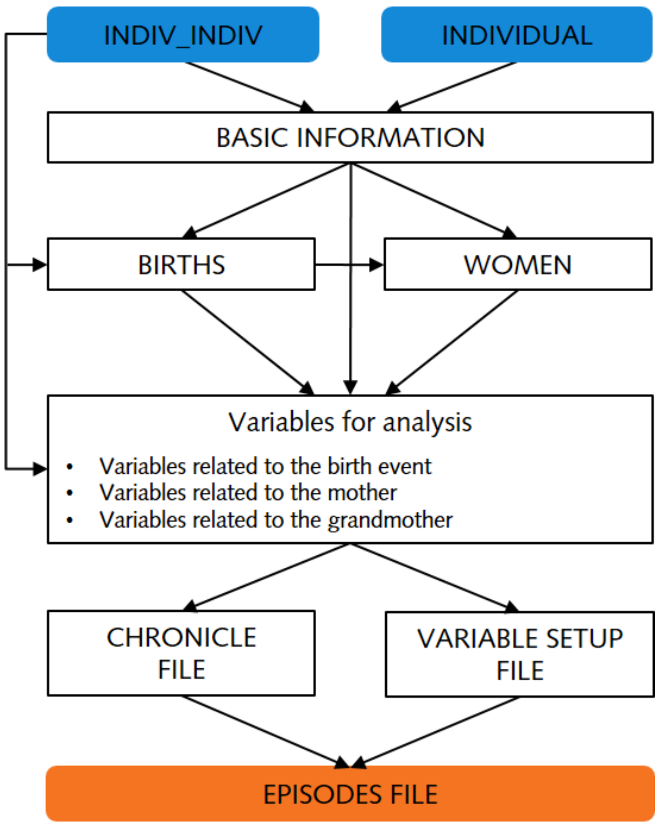
1 The "Dataset creator" and "Statistical analysis program" are attached to this article in the journal's webpage and will also be available for download in the Repository of the EHPS-net (www.ehps-net.eu).

Type on the same date. The VARIABLE SETUP FILE stores information relating to each variable included in the CHRONICLE FILE in order to facilitate the construction of the EPISODES FILE. It contains the fields Type, Transition and Duration. The field Transition is used to indicate whether a Type is a time-varying variable with Values that change at the beginning of a spell (e.g. civil status – Transition = Start), an event occurring at the end of a spell (e.g. death – Transition = End) or a time-invariant variable (e.g. sex – Transition = Invariant). The field Duration distinguishes whether the Values of a Type are valid only on their date of declaration (Duration = Instant) or between a date of declaration and the next date of declaration/End_date (Duration = Continuous). Data extractions stored using a CHRONICLE FILE and a VARIABLE SETUP FILE can be transformed into a rectangular episodes table using the program "Episodes file creator" (Quaranta, 2015, 2016).

3 DATASET CREATOR PROGRAM

The first program presented in this work (Figure 1) creates a dataset for analysing intergenerational transmissions in infant mortality along the maternal line, using as source information data stored in the IDS format in the INDIVIDUAL and INDIV_INDIV tables. Figure 2 presents the structure of this program, which is explained in detail in the current section.

Figure 2 *Structure of the "Dataset creator"*



Examples of the output produced by each step of the program are shown in a series of tables, considering data for two grandmothers, Id_I 300486006 and 400348004, drawn from the Scanian Economic Demographic Database (Bengtsson, Dribe, Quaranta, & Svensson, 2017). The INDIVIDUAL table for these women is shown in Table 1. Grandmother 300486006 had 5 children that were observed in the data, 2 of which were females and had children of their own (Id_I 990024407 and 990024527). Grandmother 400348004 also had 5 children that were observed in the data, 3 of which were females and had children of their own (Id_I 401253006, 401253005 and 401253004). The INDIV_INDIV table for these 7 women is shown in Table 2. Grandmother 300486006 had 5 grandchildren, while grandmother 400348004 had 8 grandchildren.

Table 1 *Example of an INDIVIDUAL table*

Id_I	Type	Value	Year	Month	Day
300486006	Birth_date		1818	12	25
300486006	Death_date		1887	10	4
300486006	End_observation	Death	1887	10	4
300486006	FamilyID	246618	1818	12	25
300486006	FamilyID	246618	1820	12	31
300486006	FamilyID	246618	1821	1	1
300486006	Marriage_date		1848	7	2
300486006	Marriage_date		1859	4	3
300486006	ParishID	3	1818	12	25
300486006	ParishID	3	1820	12	31
300486006	ParishID	3	1821	1	1
300486006	Sex	Female			
300486006	Start_observation	Birth	1818	12	25
300486006	Start_observation	Start_source	1821	1	1
400348004	Birth_date		1827	11	23
400348004	End_observation	Departure	1859	7	1
400348004	Marriage_date		1848	1	7
400348004	Sex	Female			
400348004	Start_observation	Birth	1827	11	23

Table 2 *Example of an INDIV_INDIV table*

	Id_I_1	Id_I_2	Relation
	300486006	990024407	Mother
	300486006	300123003	Mother
	300486006	300181004	Mother
	300486006	300181006	Mother
	300486006	990024527	Mother
Grandmothers	400348004	401253003	Mother
	400348004	401253004	Mother
	400348004	401253005	Mother
	400348004	401253007	Mother
	400348004	401253006	Mother
	990024407	500967003	Mother
	990024407	500967004	Mother
	990024407	500967005	Mother
	990024407	500967006	Mother
	990024527	990024920	Mother
	401253004	402293003	Mother
Mothers	401253004	402293004	Mother
	401253004	402293005	Mother
	401253004	402293006	Mother
	401253005	401849003	Mother
	401253006	401545003	Mother
	401253006	401545004	Mother
	401253006	401545005	Mother

From the INDIVIDUAL table the program uses the variables Birth_date, Sex, Death_date Marriage_date and, if available, Baptism, Burial, Start_observation and End_observation. From the INDIV_INDIV table the program uses the relationships Mother and Father (of a child) and Husband (of a wife).[2] No other information besides these relations and variable types are used by the program, which means that it is possible to run it to study intergenerational transmissions in infant mortality even for datasets which have not been fully transferred into IDS yet and which only contain basic information.

The program selects information from the INDIVIDUAL and INDIV_INDIV tables, conducts some simple checks for problems in the data, and makes a data extraction. The program is written modularly, with sub-programs for each part and each variable created. It gives as output a CHRONICLE FILE and a VARIABLE SETUP FILE (Quaranta, 2015), and at the end they are transformed into an episodes table for statistical analysis using the "Episodes file creator."[3]

3.1 DATA CONTROLS

Prior to running the program, it is necessary to change the working directory (i.e. the path of the folder where the IDS tables are saved) and to set the personal ado folder (i.e. the path of the folder where the IDS programs are saved). Individual identifiers in the INDIVIDUAL table (Id_I) and in the INDIV_INDIV table (Id_I_1 and Id_I_2) must be stored as numerical values.

Two types of simple checks for problems in the basic data are conducted by the program. The first of these controls whether in the INDIVIDUAL table there is more than one observation for an individual for the Types Birth_date, Death_date or Sex, as well as, if available, Baptism_date and Burial_date. The second controls whether in the INDIV_INDIV table a child is linked to more than one mother or father. If errors are found when conducting these checks a message is given. The errors must be fixed in the input tables and the program must be restarted.

3.2 BASIC INFORMATION

Using data from the INDIVIDUAL and INDIV_INDIV tables, this part of the program creates the file BASIC_INFORMATION (see Figure 2), which stores the following variables for all individuals included in the dataset: sex and the dates of birth, birth event, baptism event, death, first observation, last observation, first observed marriage, birth of first observed child and death of first observed husband. This file is used repeatedly throughout the program, to define the periods during which the individuals are under observation in the studied area and to create variables for analysis. An example of the file is shown in Table 3.

The variable Birth_event of the basic information table stores the date of birth for children who were under exposure from birth, which are those for whom a birth event was observed and recorded in the data. For children for whom the date of birth was derived indirectly from another certificate or population register (indicated by the values 'Declared' or 'Assigned' in the field Date_type), the variable Birth_event is left blank by the program. In some populations birth events are not recorded in the sources, but baptisms are recorded instead. Such information is stored in the INDIVIDUAL table using the variables Baptism or Baptism_date and assigning the value 'Event' in the field Date_type. In the BASIC_INFORMATION file the field Baptism_event is created to store the date of baptism for such cases. In this work, children whose baptism was observed are considered as having been under exposure in the study population from their date of birth.

The variable Start_observation of the BASIC_INFORMATION file defines the date when the individual is first under observation in the studied area. This variable is created by selecting for each person the first occurring date among all the events (Date_type = 'Event') or Start_observations stored in the INDIVIDUAL table. Start observations are not available in databases created using data from church books or civil registers, and in such cases the start of exposure is here defined from observed events

2 The INDIV_INDIV table stores the IDs of two individuals that are related and the type of relationship that unites them. When selecting the rows containing "Mother", "Father" and "Husband" relations, the IDs of children and wives are thus also selected and it is therefore not necessary to extract the rows containing the reciprocal relations "Child" and "Wife".

3 Version 1.1 is used. Such version of the program will be available for download in the Repository of the EHPS-net (www.ehps-net.eu).

(birth, baptism, marriage, or census registration). Similarly, the variable End_observation of the basic information table is defined as the date when the individual is for the last time under observation in the studied area. This variable is created by selecting for each person the last date among all events (Date_type = 'Event') or End_observations stored in the INDIVIDUAL table. End observations are not available in databases created using data from church books or civil registers, and the date concerning the end of exposure is instead defined by the program from observed events (marriage, death, burial, census registration or last observed birth or baptism of a child).

Table 3 *Example of the BASIC_INFORMATION file*

Id_I	**BirthDate**	**firstBday**	**Birth_event**	**Baptism_event**	**DeathDate**	**Sex**	**FirstDate**	**LastDate**	**Husband-Death-Date**	**First-Marriage-Date**	**Last-Birth-Event**
300486006	25dec1818	25dec1819	25dec1818		04oct1887	female	25dec1818	04oct1887		02jul1848	13nov1859
400348004	23nov1827	23nov1828	23nov1827			female	23nov1827	01jul1859		07jan1848	24feb1859
401253004	26mar1851	26mar1852	26mar1851			female	26mar1851	01jul1876			22jun1876
401253005	20oct1853	20oct1854	20oct1853			female	20oct1853	08jul1874			20may1873
401253006	01sep1856	01sep1857	01sep1856			female	01sep1856	06nov1890			30jan1881
401545003	27jan1877	27jan1878	27jan1877			male	27jan1877	06nov1890			
401545004	04feb1879	04feb1880	04feb1879			female	04feb1879	06nov1890			
401545005	30jan1881	30jan1882	30jan1881			female	30jan1881	06nov1890			
401849003	20may1873	20may1874	20may1873			male	20may1873	08jul1874			
402293003	15jan1870	15jan1871	15jan1870			female	15jan1870	01jul1876			
402293004	29jan1872	29jan1873	29jan1872			female	29jan1872	01jul1876			
402293005	25oct1873	25oct1874	25oct1873			male	25oct1873	01jul1876			
402293006	22jun1876	22jun1877	22jun1876			female	22jun1876	01jul1876			
500967003	18oct1881	18oct1882	18oct1881			female	18oct1881	01apr1892			
500967004	01jul1884	01jul1885	01jul1884			female	01jul1884	29oct1920		27oct1911	
500967005	18mar1888	18mar1889	18mar1888			female	18mar1888	26apri1907			
500967006	27apr1891	27apr1892	27apr1891			male	27apr1891	01apr1892			
990024407	24may1853	24may1854	24may1853			female	24may1853	01apr1892		07dec1880	27apr1891
990024527	24aug1848	24aug1849	24aug1848			female	24aug1848	23oct1879		26dec1876	24jun1877
990024920	24jun1877	24jun1878	24jun1877			male	24jun1877	23oct1879			

3.3 INFORMATION ABOUT WOMEN AND THEIR BIRTHS

Using data from the INDIV_INDIV and the BASIC_INFORMATION files, this section of the program creates the file BIRTHS (see Figure 2), which contains information for each birth given by a woman in the population. An example is shown in Table 4. The variables included in the file are: Child ID; Sex; the date of Birth, Birth event, Baptism event, First birthday, and Death; three variables with the number of total, female and male births; and three variables with the number of total, male and female infant deaths. Multiple births are recorded as a single birth event. For such cases the program registers one ChildID (the id with the lowest numerical value) and one date of death (the latter date). If they are born one day apart, the first of these dates of births is recorded. The variables total, male and female births and infant deaths have values 0 or 1 for singleton births but can have values above 1 for multiple births. Some of the variables included in the file are left blank for some individuals. For example, as explained in the previous section, the date of birth event has values only for children for whom the actual birth was observed in the population during the study period and was recorded in the data. As a final step, the indicator variable missingbdate is also added to the BIRTHS file, which indicates whether the date of birth of any of the woman's children is missing. Such variables allow selecting cases for which the variables for analysis are correctly defined in Part 4 of the program. In particular, to identify birth order correctly, the dates of births of all children born to a woman must be available.

Table 4 *Example of the BIRTHS file*

MotherID	BirthDate	firstBday	Birth_event	Baptism_event	Sex	nbirths	ChildID	DeathDate	LastChild-Date	border	Female_Birth	Male_Birth	Female_Infant-Death	Male_Infant-Death	Infant-Death	missing-bdate
300486006	24aug1848	24aug1849	24aug1848		female	1	990024527		23oct1879	1	1	0	0	0	0	0
300486006	27jun1850	27jun1851	27jun1850		female	1	300181004	31mar1857	31mar1857	2	1	0	0	0	0	0
300486006	24may1853	24may1854	24may1853		female	1	990024407		01apr1892	3	1	0	0	0	0	0
300486006	19jun1856	19jun1857	19jun1856		male	1	300181006	17aug1856	17aug1856	4	0	1	0	1	1	0
300486006	13nov1859	13nov1860	13nov1859		female	1	300123003		22feb1892	5	1	0	0	0	0	0
400348004	21apr1849	21apr1850	21apr1849		female	1	401253003		01jul1859	1	1	0	0	0	0	0
400348004	26mar1851	26mar1852	26mar1851		female	1	401253004		01jul1876	2	1	0	0	0	0	0
400348004	20oct1853	20oct1854	20oct1853		female	1	401253005		08jul1874	3	1	0	0	0	0	0
400348004	01sep1856	01sep1857	01sep1856		female	1	401253006		06nov1890	4	1	0	0	0	0	0
400348004	24feb1859	24feb1860	24feb1859		male	1	401253007		01jul1859	5	0	1	0	0	0	0
401253004	15jan1870	15jan1871	15jan1870		female	1	402293003		01jul1876	1	1	0	0	0	0	0
401253004	29jan1872	29jan1873	29jan1872		female	1	402293004		01jul1876	2	1	0	0	0	0	0
401253004	25oct1873	25oct1874	25oct1873		male	1	402293005		01jul1876	3	0	1	0	0	0	0
401253004	22jun1876	22jun1877	22jun1876		female	1	402293006		01jul1876	4	1	0	0	0	0	0
401253005	20may1873	20may1874	20may1873		male	1	401849003		08jul1874	1	0	1	0	0	0	0
401253006	27jan1877	27jan1878	27jan1877		male	1	401545003		06nov1890	1	0	1	0	0	0	0
401253006	04feb1879	04feb1880	04feb1879		female	1	401545004		06nov1890	2	1	0	0	0	0	0
401253006	30jan1881	30jan1882	30jan1881		female	1	401545005		06nov1890	3	1	0	0	0	0	0
990024407	18oct1881	18oct1882	18oct1881		female	1	500967003		01apr1892	1	1	0	0	0	0	0
990024407	01jul1884	01jul1885	01jul1884		female	1	500967004		29oct1920	2	1	0	0	0	0	0
990024407	18mar1888	18mar1889	18mar1888		female	1	500967005		27apr1907	3	1	0	0	0	0	0
990024407	27apr1891	27apr1892	27apr1891		male	1	500967006		01apr1892	4	0	1	0	0	0	0
990024527	24jun1877	24jun1878	24jun1877		male	1	990024920		23oct1879	1	0	1	0	0	0	0

The file WOMEN, containing information on exposure for each woman is also created. An example is shown in Table 5. Each unique mother is selected from the file BIRTHS produced above and based on the data stored in the file BASIC_INFORMATION three indicator variables are generated to measure (1) whether the woman was observed until her 50th birthday, (2) whether the woman was observed until her 50th birthday or her death date, and (3) whether the woman and her first husband were observed until the woman's 50th birthday. The purpose of these three indicator variables is to identify the subset of women that were under observation during their entire reproductive periods, allowing to test different specifications of that subset.

Table 5 *Example of the WOMEN file*

Id_I	BirthDate	firstBday	Birth_event	Baptism_event	DeathDate	FirstDate	LastDate	Husband-Death-Date	First-Marriage-Date	Last-Birth-Event	first_child_bdate	Obs_2050	Obs_2050-death	Obs_2050-hb
300486006	25dec1818	25dec1819	25dec1818		04oct1887	25dec1818	04oct1887		02jul1848	13nov1859	24aug1848	1	1	1
400348004	23nov1827	23nov1828	23nov1827			23nov1827	01jul1859		07jan1848	24feb1859	21apr1849	0	0	0
401253004	26mar1851	26mar1852	26mar1851			26mar1851	01jul1876			22jun1876	15jan1870	0	0	0
401253005	20oct1853	20oct1854	20oct1853			20oct1853	08jul1874			20may1873	20may1873	0	0	0
401253006	01sep1856	01sep1857	01sep1856			01sep1856	06nov1890			30jan1881	27jan1877	0	0	0
990024407	24may1853	24may1854	24may1853			24may1853	01apr1892		07dec1880	27apr1891	18oct1881	0	0	0
990024527	24aug1848	24aug1849	24aug1848			24aug1848	23oct1879		26dec1876	24jun1877	24jun1877	0	0	0

3.4 VARIABLES FOR ANALYSIS

Part 4 of the program has the purpose of generating the variables to be included in the empirical analysis, using the information contained in the files INDIV_INDIV, BASIC_INFORMATION, BIRTHS and WOMEN, as is shown in Figure 2. The variables are constructed in a modular basis, with specific subprograms for each of them. Each is stored in a separate file, having the same format as the CHRONICLE FILE (ChildID, Type, Value and a time stamp if applicable). The following variables are created:

Variables related to the birth event (third generation):

- ChildID: Identifier of the child.
- Child_birthdate: birth date of the child.
- Birth_order: observed birth order of the child. The effective birth order cannot be identified for all children since not all mothers are under observation for the entire reproductive period.
- Child_sex: sex of child (female, male, or multiple in the case of multiple births of different sexes).
- N_births: number of births at the specific birth event (1 for singleton birth and 2 or above for multiple births).
- Infant_death: infant death event for the child. It gets value '1' if the child dies before the first birthday. The date of death is assigned to the time stamp. This variable represents the event of interest in the empirical analysis.
- At_risk: time-varying variable identifying whether a child is at risk of dying in infancy. It is created only for children for whom a birth or a baptism event was observed, based on the variables Birth_event and Baptism_event of the file BASIC_INFORMATION. Two rows are created for each individual. The first row gets value '1', with the time stamp of the child's date of birth. The second row gets value '0', with the time stamp of the date of whichever occurs first, the first birthday, death or out-migration (if observed).
- Under observation: time-varying variable used to define the periods during which a child is under observation and below age 1. In the main statistical models children are considered under exposure (i.e. the variable At_risk created above has value '1') until their first birthday unless they die or out-migrate before then. However, given that information on migration is not available in databases created using church books or civil registers, when using such sources the size of the population at risk may be overestimated. To test for this, one of the models in the sensitivity analysis only considers rows with value '1' for the variable Under_observation. All children for whom a birth event was observed are assigned value '1' on their date of birth and value '0' on their last date of observation, based on the following criteria. Children who died before their first birthday are given as last date of observation their date of death. Children for whom there is additional observations in the data – the date of the birth of successive offspring for the mother, the mother's death, the child's death (after age 1), the child's marriage, the birth of the child's own children, or a date of out-migration for either the mother or the child – are given as last date of observation the date of their first birthday. Children for whom none of these dates are available are given as last date of observation their date of birth, and therefore contribute with only one day to the population at risk.

Variables related to the mother (second generation):

- MotherID: Identifier of the mother.
- Mother_birthdate: birth date of the mother.

Variables related to the grandmother (first generation):

- GrandmotherID: Identifier of the grandmother.
- Grandmother_birthdate: birth date of the grandmother.
- G0_Births/G0_F_Births/G0_M_Births: number of observed births (total/females/males) for the grandmother. This variable measures the observed rather than the effective number of

births, since not all grandmothers are under observation in the studied area for their entire reproductive period.

- G0_Births_unknownBdate: number of observed children with unknown birth dates for the grandmother.
- G0_mult_b: number of observed multiple birth events for the grandmother.
- G0_InfD/G0_F_InfD/G0_M_InfD: number of observed infant deaths (total/females/ males) for the grandmother. This variable measures the observed rather than the effective number of infant deaths, since not all grandmothers are under observation in the studied area for their entire reproductive period.
- GO_under_age_50: time-varying variable considering a classification of the age of the grandmother. It gets value '1' if the grandmother's age is less than 50, value '0' if her age is equal to or above 50, and value '-1' if her date of birth is unknown.
- G0_obs_2050: variable that has value '1' if the grandmother was under observation in the study area at least from age 20 to age 50.
- G0_obs_2050death: variable that has value '1' if the grandmother was under observation in the study area at least from age 20 to age 50 or died between these age ranges.
- G0_obs_2050husband: variable that has value '1' if the grandmother was under observation in the study area at least from age 20 to age 50, and her first observed husband (i.e. the grandfather) was under observation during this age range.

3.5 EPISODES FILE

As discussed in section 2, data extractions should be included in a CHRONICLE FILE, which stores variables for analysis and a VARIABLE SETUP FILE, which stores information relating to each variable. Using the program "Episodes file creator" (Quaranta, 2016) the information contained in these files is transformed into a rectangular file, the so-called episodes table, which is ready to be used for statistical analysis.

In part 4 of the current program the variables for analysis were created on a modular basis, saving each variable in a separate file. The CHRONICLE FILE is created by the program by simply appending together the information contained in these separate files. An example of part of this file is shown in Table 6. Only three variables are included in the example, one relating to the birth event (birth order), one to the mother (birthdate) and one to the grandmother (number of infant deaths). It is also possible to append to the CHRONICLE FILE variables created by the researcher using other programs, as long as such variables comply with the specifications of these files (for such specifications see Quaranta, 2015).

Table 6 *Example of the CHRONICLE file*

ChildID	Year	Month	Day	Type	Value
402293003				Birth_order	1
402293004				Birth_order	2
402293005				Birth_order	3
402293006				Birth_order	4
401849003				Birth_order	1
401545003				Birth_order	1
401545004				Birth_order	2
401545005				Birth_order	3
500967003				Birth_order	1
500967004				Birth_order	2
500967005				Birth_order	3
500967006				Birth_order	4
990024920				Birth_order	1

Table 6 continued on next page

ChildID	Year	Month	Day	Type	Value
402293003	1851	3	26	Mother_birthdate	
402293004	1851	3	26	Mother_birthdate	
402293005	1851	3	26	Mother_birthdate	
402293006	1851	3	26	Mother_birthdate	
401849003	1853	10	20	Mother_birthdate	
401545003	1856	9	1	Mother_birthdate	
401545004	1856	9	1	Mother_birthdate	
401545005	1856	9	1	Mother_birthdate	
500967003	1853	5	24	Mother_birthdate	
500967004	1853	5	24	Mother_birthdate	
500967005	1853	5	24	Mother_birthdate	
500967006	1853	5	24	Mother_birthdate	
990024920	1848	8	24	Mother_birthdate	
401545003				G0_InfD	0
401545004				G0_InfD	0
401545005				G0_InfD	0
401849003				G0_InfD	0
402293003				G0_InfD	0
402293004				G0_InfD	0
402293005				G0_InfD	0
402293006				G0_InfD	0
500967003				G0_InfD	1
500967004				G0_InfD	1
500967005				G0_InfD	1
500967006				G0_InfD	1
990024920				G0_InfD	1

The *VARIABLE SETUP FILE* is also created. The field Transition is set to 'Invariant' for the variables Birth_order, N_births, Child / Mother / Grandmother_birthdate, MotherID, GrandmotherID, Child_sex and all the variables relating to the grandmother, except G_under_age_50. It is set to 'Start' for the variables Under_observation, G_under_age_50 and At_risk, and to 'End' for the variable Infant_death, which is the outcome of interest of the study. An example of the VARIABLE SETUP FILE is shown in Table 7.

Table 7 *Example of the VARIABLE SETUP file*

Type	Transition	Duration
Child_sex	Invariant	Continuous
Birth_order	Invariant	Continuous
N_births	Invariant	Continuous
Infant_death	End	Continuous
Child_birthdate	Invariant	Continuous
Mother_birthdate	Invariant	Continuous
MotherID	Invariant	Continuous
Under_observation	Start	Continuous
At_risk	Start	Continuous
GrandmotherID	Invariant	Continuous
G0_Births_unknownBdate	Invariant	Continuous
G0_F_Births	Invariant	Continuous

Table 7 continued on next page

Type	Transition	Duration
G0_M_Births	Invariant	Continuous
G0_M_InfD	Invariant	Continuous
G0_F_InfD	Invariant	Continuous
G0_InfD	Invariant	Continuous
G0_Births	Invariant	Continuous
G0_mult_b	Invariant	Continuous
G0_obs_2050husband	Invariant	Continuous
G0_obs_2050death	Invariant	Continuous
G0_obs_2050	Invariant	Continuous
Grandmother_birthdate	Invariant	Continuous
GO_under_age_50	Start	Continuous
G1_obs_2050husband	Invariant	Continuous

The information stored in the CHRONICLE FILE and VARIABLE SETUP FILE is finally transformed into an EPISODES FILE, using the program "Episodes file creator". The program deletes from the table all rows where the individual is not at risk of experiencing the event of interest (in this case all periods when a child is not at risk of dying in infancy, i.e. At_risk = 0) and deletes the variable At_risk after removing such rows. All other variables included in the CHRONICLE FILE become the columns of the EPISODES FILE. Missing values of numerical variables are replaced with '-1'. The variables date1 and date2 are added. Date1 is the start date of each spell, which in this case is equivalent to the date of birth of the child or to the date of change in any time-varying covariate. Date2 is the end date of each spell which in this case is equivalent to the date of failure (infant death) or censoring (date of first birthday or out-migration) or to the date of change in any time-varying covariate. An example of the EPISODES FILE is shown in Table 8.

3.6 ERASE TEMPORARY FILES

Throughout the program various temporary files are created. These files are no longer needed for analysis and are therefore erased by the program.

4 STATISTICAL ANALYSIS PROGRAM

The current section discusses in more detail the program developed to run the statistical analysis. As is shown in Figure 3, the program uses the EPISODES FILE produced by the previously discussed programs and stores the results of the analysis. The program is divided into different sections. In part 1, additional variables are constructed. In part 2, the study sample is selected. In part 3, a graph of the infant mortality rate over the study period is created, showing calculations for the entire population as well as the study sample. In part 4, some tests of the quality of the data in the study sample are made. In part 5, descriptive statistics are produced, and in part 6, model estimations are calculated, producing Excel files with the respective outputs.

Figure 3 *Structure of the "Statistical analysis program"*

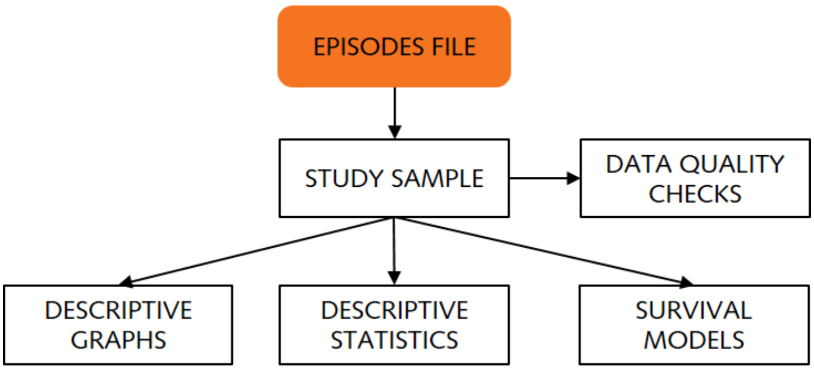


Table 8 *Example of the EPISODES file*

Id_I	date1	date2	GO_ under_ age_ 50	Under_ obser- vation	Birth_ order	Child_ sex	G0_ Births	G0_ Births_ unknown- Bdate	G0_F_ Births	G0_F_ InfD	G0_ InfD	G0_M_ Births	G0_M_ InfD	G0_ mult_b	G0_ obs_ 2050	G0_ obs_ 2050- death	G0_ obs_ 2050- husband	G1_ obs_ 2050- husband	Grand- motherID	MotherID	N_ births	Infant_ death	Child_ birthdate	Grand_ mother_ birthdate	Mother_ birthdate
401545003	27jan1877	23nov1877	1	1	1	male	5	0	4	0	0	1	0	0	0	0	0	0	400348004	401253006	1		27jan1877	23nov1827	01sep1856
401545003	23nov1877	27jan1878	0	1	1	male	5	0	4	0	0	1	0	0	0	0	0	0	400348004	401253006	1		27jan1877	23nov1827	01sep1856
401545004	04feb1879	04feb1880	0	1	2	female	5	0	4	0	0	1	0	0	0	0	0	0	400348004	401253006	1		04feb1879	23nov1827	01sep1856
401545005	30jan1881	30jan1882	0	1	3	female	5	0	4	0	0	1	0	0	0	0	0	0	400348004	401253006	1		30jan1881	23nov1827	01sep1856
401849003	20may1873	20may1874	1	1	1	male	5	0	4	0	0	1	0	0	0	0	0	0	400348004	401253005	1		20may1873	23nov1827	20oct1853
402293003	15jan1870	15jan1871	1	1	1	female	5	0	4	0	0	1	0	0	0	0	0	0	400348004	401253004	1		15jan1870	23nov1827	26mar1851
402293004	29jan1872	29jan1873	1	1	2	female	5	0	4	0	0	1	0	0	0	0	0	0	400348004	401253004	1		29jan1872	23nov1827	26mar1851
402293005	25oct1873	25oct1874	1	1	3	male	5	0	4	0	0	1	0	0	0	0	0	0	400348004	401253004	1		25oct1873	23nov1827	26mar1851
402293006	22jun1876	01jul1876	1	1	4	female	5	0	4	0	0	1	0	0	0	0	0	0	400348004	401253004	1		22jun1876	23nov1827	26mar1851
500967003	18oct1881	18oct1882	0	1	1	female	5	0	4	0	1	1	1	0	1	1	1	0	300486006	990024407	1		18oct1881	25dec1818	24may1853
500967004	01jul1884	01jul1885	0	1	2	female	5	0	4	0	1	1	1	0	1	1	1	0	300486006	990024407	1		01jul1884	25dec1818	24may1853
500967005	18mar1888	18mar1889	0	1	3	female	5	0	4	0	1	1	1	0	1	1	1	0	300486006	990024407	1		18mar1888	25dec1818	24may1853
500967006	27apr1891	01apr1892	0	1	4	male	5	0	4	0	1	1	1	0	1	1	1	0	300486006	990024407	1		27apr1891	25dec1818	24may1853
990024920	24jun1877	24jun1878	0	1	1	male	5	0	4	0	1	1	1	0	1	1	1	0	300486006	990024527	1		24jun1877	25dec1818	24aug1848

4.1 CATEGORIZATION OF VARIABLES

Additional variables to be used in the statistical models and for refinements of the study sample are created by extracting or categorizing information from the variables included in the EPISODES FILE.

- Child_birthyear: year of birth of the child. This variable is included in the basic models and it can also be used to obtain descriptive information about the study sample and to make further refinements to it.
- Child_birthdate_cent: birthdate of the child. The variable is centred by subtracting from each date of birth January 1st of the average year of birth of all children included in the sample. Centring this variable does not affect the estimations, but it facilitates the interpretation of the results.
- Birth_order_cat: categorical variable of the birth order of the child: 1, 2, 3, 4, 5, 6, 7+ (7 or more). The reference category in the statistical models is set at 1.
- Mother_birthyear: year of birth of the mother. This variable can be used for descriptive information about the sample and to make further refinements to it.
- Grandmother_birthyear: year of birth of the grandmother. This variable can be used for descriptive information about the sample and to make further refinements to it.
- G0_InfD_cat: categorical variable of the number of observed infant deaths for the grandmother: 0, 1, 2+ (2 or more). The reference category for the statistical models is set at 0.
- G0_Births_cat: categorical variable of the number of observed births for the grandmother: 2, 3, 4-6, 7+ (7 or more). The reference category for the statistical models is set at 2.
- Mother_age: continuous variable of the age of mother at the birth of the child.
- Mother_age_cat: categorical variable of age of mother at the birth of the child, classified into three groups: 15-24, 25-34, 35-50. The reference category for the statistical models is set at 25-34.

4.2 SELECTION OF THE STUDY SAMPLE

The sample selected for analysis consists of women for whom information is available about the births and infant mortality/survival of her offspring and about the infant mortality/survival of their mothers (the 'grandmothers'). The study focuses only on singleton births. In order to be able to calculate an intergenerational transmission variable, only cases where the grandmother gave birth to at least two children are selected. To exclude possible errors in the data, cases with unknown date of birth for the mother, where the mother was linked to any child born before her 15th birthday or after her 50th birthday, or where the grandmother was linked to any grandchildren whose birth occurred before her 30th birthday or after her 100th birthday are deleted from the study sample. Children with unknown sex are also dropped.

Descriptive statistics are calculated and shown for the year of birth of the children, mothers and grandmothers. These statistics allow the researcher to make further refinements to the study sample if desired, by, for example, restricting the dataset to children/mothers/grandmothers born before or after a certain year.

4.3 GRAPH OF INFANT MORTALITY RATE

To provide additional background information for the area and period under study, infant mortality rates are calculated and a graph is produced. The rate is first calculated for the entire population directly from the source material (i.e. the INDIVIDUAL table), and therefore considering all individuals born in the area during the study period. The rate is also calculated for the study sample using the dataset created for analysis, and therefore considering all children (third generation) born in the area and study period and whose maternal grandmothers gave birth to at least two children. A graph containing the lines of the two rates is produced. There are some specifications which can be set by the researcher to change the axis of the graphs, depending on the data and period considered.

4.4 DATA QUALITY CHECKS OF THE STUDY SAMPLE

Part 4 of the program is dedicated to produce statistics and graphs which allow the researcher to assess the quality of the data and whether the study sample may have problems of selection with respect to the population as a whole. For such purpose, the percentage of births and infant deaths by sex and the percentage of infant deaths taking place in the first month of life are calculated and shown. Cumulative hazard curves of the hazard of death within the first year of life are made. All calculations and graphs are shown for the study sample as well as the entire population in order to provide a point of comparison.

4.5 DESCRIPTIVE STATISTICS

In this part of the program statistics are generated to describe the variables included in the statistical models. The mean, standard deviation and minimum and maxim values are shown for the continuous variables Child_birthyear and Child_birthday_cent, and the percentage distribution is shown for the categorical variables Child_sex, G0_InfD_cat, G0_Births_cat, Birth_order_cat and Mother_age_cat. The descriptive statics are shown on the results screen, but they are also saved to an Excel file (Descriptive_statistics.xlsx).

4.6 SURVIVAL MODELS

The final part of the program is devoted to the statistical analysis. The data is set for survival analysis, considering the age of the child as the time variable. Each child is followed from the date of birth until the date of the first birthday, or until outmigration or death if they happen before age 1. Graphs with hazard curves are first produced, estimating different curves by number of infant deaths of the maternal grandmother, by sex of the child and by mother's age.

Three different survival models are estimated: a Cox model, a Weibull model with shared frailty on the mother and a Weibull model with shared frailty on the grandmother. To measure intergenerational transfers in infant mortality, the main explanatory variable included in the model is the number of observed infant deaths for the grandmother. The models also control for the number of observed births for the grandmother, the sex of child, the birth order of child, the birth date of the child and the age of the mother. The results of the models are shown on the screen and are also saved into an Excel file (Results_survival.xlsx). Tests to check whether the proportional hazards assumption is violated are made after the Cox model.

The sensitivity analysis tests whether changes in the selection of the data affect the findings of the study. The results are saved into an Excel file (Results_survival_sensitivity.xlsx). Six Weibull models with shared frailty on the mother are estimated for the following subsamples:

- Model 1: Children under observation (see chapter 3, part 4).
- Model 2: Grandmothers observed at least between ages 20-50.
- Model 3: Grandmothers observed at least between ages 20-50 or death.
- Model 4: Grandmothers observed at least between ages 20-50 and grandfathers observed at least until the grandmothers' 50th birthday.
- Model 5: Grandmothers observed at least between ages 20-50 and no unknown dates of birth for any of her children.
- Model 6: Children under observation (see chapter 3, part 4) and grandmothers observed at least between ages 20-50.

5 CONCLUSIONS

The literature on intergenerational transmissions in infant mortality is rather limited. A comparative project across populations in Belgium (Donrovich, Puschmann, & Matthijs 2018), the Netherlands (Van Dijk & Mandemakers, 2018), Norway (Sommerseth, 2018), Northern Sweden (Broström, Edvinsson,

& Engberg, 2018) and Southern Sweden (Quaranta, 2018) was started in order to expand such literature as well as to show the advantages of using the IDS for research in historical demography. The current article presented two programs written for STATA, which allow studying whether the likelihood of death in infancy is influenced by the number of infant deaths experienced by the maternal grandmother. The programs use input data stored in the IDS tables to construct the dataset for analysis and run statistical models, producing output Excel files with descriptive statistics and model results. They were developed modularly, to simplify making changes or additions to the code. The programs can be used with longitudinal demographic datasets containing data obtained from church books, civil registers, or populations registers, as long as the data has been transferred to the IDS.

This work forms part of the first project that uses the same programs to conduct fully comparable research across more databases. One of the main challenges faced when developing the programs was making them compatible for databases created from either church books/civil registers or population registers. Although in these types of databases variables are stored in the same way in the IDS tables, the main distinction across them is how individuals under exposure are defined and which type of information from the IDS tables is considered to make such definition. Except for such challenge, the simple and standardized nature of the IDS tables (Alter & Mandemakers, 2014; Alter, Mandemakers, & Gutmann, 2009) and of the solutions previously developed to make data extractions (Quaranta, 2015) have facilitated creating these programs. Thanks to such IDS developments, the amount of work needed for writing generic code that was compatible for any database and the difficulties encountered were not much larger than when writing code to be used only by one specific database. Although there is some investment involved in transferring databases into the IDS, this article and the project which it forms part of show that when using the IDS it is possible to write common software that allows producing fully comparable, transparent, and replicable studies.

ACKNOWLEDGEMENTS

Funding from the Jan Wallander and Tom Hedelius Foundation is gratefully acknowledged.

This project has received funding from the European Union's Horizon 2020 research and innovation programme under the Marie Sklodowska-Curie grant agreement No 676060

This article has been written in the context of the LONGPOP project; this project has received funding from the European Union's Horizon 2020 research and innovation programme under the Marie Sklodowska-Curie grant agreement No 676060. This publication reflects only the author's view and the Research Executive Agency is not responsible for any use that may be made of the information it contains.

REFERENCES

Alter, G., & Mandemakers, K. (2014). The Intermediate Data Structure (IDS) for longitudinal historical microdata, version 4. *Historical Life Course Studies, 1*, 1-26. Retrieved from http://hdl.handle.net/10622/23526343-2014-0001?locatt=view:master

Alter, G., Mandemakers, K., & Gutmann, M. (2009). Defining and distributing longitudinal historical data in a general way through an Intermediate Structure. *Historical Social Research, 34*(3), 78-114. doi: 10.12759/hsr.34.2009.3.78-114

Bengtsson, T., Dribe, M., Quaranta, L., & Svensson, P. (2017). *The Scanian Economic Demographic Database. Version 5.1 (machine-readable database)*. Lund: Lund University, Centre for Economic Demography.

Broström, G., Edvinsson, S., & Engberg, E. (2018). Intergenerational transfers of infant mortality in 19th century northern Sweden. *Historical Life Course Studies, 7*, 106-122. Retrieved from http://hdl.handle.net/10622/23526343-2018-0005?locatt=view:master

Cox, D. (1972). Regression models and life-tables. *Journal of the Royal Statistical Society. Series B (Methodological), 34*(2), 187-220. Retrieved from http://www.jstor.org/stable/2985181

Das Gupta, M. (1990). Death clustering, mother's education and the determinants of child mortality in rural Punjab, India. *Population Studies, 44*, 489-505. doi: 10.1080/0032472031000144866

Donrovich, R., Puschmann, P., & Matthijs, K. (2018). Mortality clustering in the family: Fast life history trajectories and the intergenerational transfer of infant death in late 19th- and early 20th-century Antwerp, Belgium. *Historical Life Course Studies*, *7*, 47-68. Retrieved from http://hdl.handle.net/10622/23526343-2018-0006?locatt=view:master

Edvinsson, S., Brändström, A., Rogers, J., & Broström, G. (2005). High-risk families: The unequal distribution of infant mortality in nineteenth-century Sweden. *Population Studies, 59*(3), 321-337. doi: 10.1080/00324720500223344

Janssens, A., Messelink, M., & Need, A. (2010). Faulty genes or faulty parents? Gender, family and survival in early and late childhood in the Netherlands, 1860-1900. *The History of the Family, 15*, 91-108. doi: 10.1016/j.hisfam.2010.01.005

Quaranta, L. (2015). Using the Intermediate Data Structure (IDS) to construct files for statistical analysis. *Historical Life Course Studies, 2*, 86-107. Retrieved from http://hdl.handle.net/10622/23526343-2015-0007?locatt=view:master

Quaranta, L. (2016). STATA programs for using the Intermediate Data Structure (IDS) to construct files for statistical analysis. *Historical Life Course Studies, 3*, 1-19. Retrieved from http://hdl.handle.net/10622/23526343-2016-0001?locatt=view:master

Quaranta, L. (2018). Intergenerational transfers in infant mortality in southern Sweden, 1740-1968. *Historical Life Course Studies, 7*, 88-105. Retrieved from http://hdl.handle.net/10622/23526343-2018-00013?locatt=view:master

Quaranta, L., Broström, G., Van Dijk, I., Donrovich, R., Edvinsson, S., Engberg, E., Mandemakers, K., Matthijs, K., Puschmann, P., & Sommerseth, H. (2017). Intergenerational transfers of infant mortality in historical contexts: a comparative study of five European populations. Paper presented at the SSHA Conference, Chicago.

Sommerseth, H. L. (2018). Intergenerational transfer in infant mortality in northern Norway during the 19th and early 20th century. *Historical Life Course Studies, 7*, 69-87. Retrieved from http://hdl.handle.net/10622/23526343-2018-0008?locatt=view:master

Stead, W., Hammond, W., & Straube, M. (1982). A chartless record – is it adequate? *Proceedings of the Annual Symposium on Computer Application in Medical Care,* 89-94.

Vandezande, M. (2012). *Born to die. Death clustering and the intergenerational transmission of infant mortality, the Antwerp District, 1846-1905*. Unpublished Ph.D. thesis at University of Leuven.

Van Dijk, I., & Mandemakers, K. (2018). Like Mother, like Daughter. Intergenerational transmission of infant mortality clustering in Zeeland, the Netherlands, 1833-1912. *Historical Life Course Studies, 7*, 28-46. Retrieved from http://hdl.handle.net/10622/23526343-2018-0003?locatt=view:master