

SURVEY OF HISTORICAL DATABASES WITH LONGITUDINAL MICRO-DATA

For more information about this questionnaire or questions about entering specific information, please mail George Alter (alter@indiana.edu) and/or Kees Mandemakers (kma@iisg.nl).

1/ Identifying information

	Name of database:	BALSAC population register
	Location:	University of Quebec at Chicoutimi, Quebec, Canada
	Web-address:	http://www.uqac.ca/balsac/
	Name contact person:	Gerard Bouchard
	Email-address contact person:	gerard_bouchard@uqac.ca
	Postal Address:	Projet BALSAC Universite du Quebec a Chicoutimi 555 boul. de l'Universite Chicoutimi, Quebec Canada G7H 2B1
	Tel. number:	(418)-545-5517

2/ Main objective and scope of the database:

Given its particular history and its current structure, the population of Quebec meets various conditions that make it particularly propitious to the construction of a historical, genealogical, and demographic database. One of these conditions is the founder effect resulting from the immigration pattern from the beginning of the 17th century up to 1760. Most immigrants who have settled in the Saint-Laurent valley during this period came from France. As a consequence, a very large proportion of the contemporary Francophone population in Quebec carries genes inherited from these pioneers. From 1820-1830 onwards, under the pressure of high fertility and land scarcity, the settlement expanded beyond the Saint-Laurence Valley so that the 19th century witnessed the creation of a dozen regions or so. The contemporary structure of the population is also accounted for by the migration streams that took place over the years. After 1760, New France became a British colony and, as a result, a strong non-Francophone immigration ensued. However, most of these immigrants and their descendants settled in the western part of the province. Therefore, the present-day 7 million population is stratified into a Francophone segment inherited from the initial stock, and a multi-ethnic component located mostly in the Montreal region. Because of this structural diversity, the Quebec population lends itself to fascinating research projects. It is also the perfect setting for the creation of a genealogical population register which aims at recreating the diachronic trace of individuals and families over nearly four centuries. Another positive factor consists in the very high quality of the Quebec vital data registration system.

The BALSAC population register is a computerized database designed for the automatic construction of family histories as well as ascending and descending genealogies. The information contained in the database comes primarily from marriage records. The project was initiated in 1972 at the University of Quebec at Chicoutimi and the overall, long term objective is to extend the register to the entire population of Quebec from the onset of the settlement (at the beginning of the 17th century) up to recent years. As of now, the data entry is completed up to 1939 for the entire province ⁽¹⁾, the linkage work is completed up to 1900 and both tasks are ongoing. For the Saguenay, a northeastern region of Quebec, the BALSAC database also contains birth and death records of the entire population (1838 to 1971).

The BALSAC register is widely used in the field of human genetics (population genetics, genetic epidemiology). It also supports a number of research pertaining to historical and social sciences.

⁽¹⁾ For the period prior to 1800, the marriages in BALSAC were provided by PRDH, University of Montreal.

3/ Sources: Please enter Yes or No and the time period for the main sources included in the database

Yes /no	Start year	End year	Type of source	Comments
Y	1838	1985	Baptisms	For the Saguenay region
Y	1621	1993	Marriages from church registers	For the entire Quebec province (completed up to 1939)
Y	1839	1978	Burials	For the Saguenay region

N			Population registers (continuous) maintained by a church	
N			Civil birth certificates	
Y	1969	1993	Civil marriage certificates	Partly processed
N			Civil death certificates	
N			Civil population registers	
N			Census	
N			Nominative lists	
Y/N			Other:	
Y/N			Other:	
Y/N			Other:	
Y/N			Other:	

4/ How was the sample (or samples) defined?

The entire Quebec population is recorded through the marriage certificates from church registers.

For the Saguenay region, all the vital data (births, marriages, deaths) have been used.

5/ Units of observation:

Please enter Y or N for each unit that can be followed over time

		Comments:
Y	Individuals	
Y	Married couples	
Y	Families	
N	Households	
N	Farms	
Y/N	Other:	
Y/N	Other:	
Y/N	Other:	

6/ Describe the geographic area under observation

The entire province of Quebec, Canada.

7/ Is information available about related individuals who are not in the sample?

Examples:

Marriage registers often include occupations of parents.

A population register sample may include everyone in the household of an individual in the sample.

Yes. There is information about individuals mentioned in the vital records but never present in the Quebec territory (Example: an immigrant's parents at his marriage).

8/ What events can be identified? Do events have dates?

Y/N	Event	Are these events dated? Y/N/P(= partial dates, e.g. year only)	
Y	Birth	Y	Comment: For the Saguenay region only
Y	Marriage	Y	Comment: For the entire province
Y	Death	Y	Comment: For the Saguenay region only
N	Migration	Y/N/P	Comment:
N	Other:	Y/N/P	Comment:
N	Other:	Y/N/P	Comment:
N	Other:	Y/N/P	Comment:

9/ These questions describe the way observation is censored.

A) How do individuals enter observation?

Individuals enter observation through the first civil record where they appear as subject or as parent.

B) How do individuals leave observation?

Individuals leave observation through the last civil record where they appear as subject or as parent.

C) Are some entry or exit dates unknown?		
No. All entries have dates.		
D) Are some entries or exits interval censored (i.e. the exact date is unknown, but it can be located between two known dates)?		
No. All dates are exact.		
10/ Residence and Household (Y/N/Partly)		
Y	Can observations be linked to residential locations?	
N	Are the dates and locations of movements within the observation area recorded?	
N	Are all individuals who lived in the households of members of the sample recorded?	
11/ Kinship relations		
A) How is kinship recorded in the sources?		
The database structure itself allows for the reconstruction of kinship links among all individuals. Moreover, canonical consanguinity for spouses declared in the marriage certificates is recorded.		
B) How deep (number of generations) is the available kinship information?		
The database structure and depth allow for the reconstruction of kinship links up to 17-18 generations (10 generations on average).		
12/ Linkage		
Which sources and units of observation have been linked?		
	Y/N/Partly	Comments:
Births/Baptisms	Partly	For the Saguenay region only
Marriages	Y	
Deaths/Burials	Partly	For the Saguenay region only
Population registers	Y/N/P	
Census	Y/N/P	

Nominative lists	Y/N/P	
Other:	Y/N/P	
Other:	Y/N/P	
Other:	Y/N/P	
Other:	Y/N/P	

How is linkage represented in the database? For example, do all occurrences of an individual include a universal identification number? Are records linked to each other but not to a universal ID?

There is a universal identification number for each individual, each couple and each family.

13 / What data structures have been added to the information in the sources?

		Comments:
Y	Date of entry and date of exit by individual	
Y	Events by individual	
Y	Time constant information (date of birth, sex, etc.) by individual	
Y	Husband-Wife pairs	
Y	Mother-child and Father-child pairs	
Y/N	Other:	
Y/N	Other:	
Y/N	Other:	
Y/N	Other:	

14/ What reference/coding systems have been linked to the data?

Y	Occupational titles (like HISCO): Home made dictionary.
Y	Locations (including geo-referenced systems): Home made dictionary.
Y	Other (religion, civil status etc.): Home made dictionary for civil status.
Y/N	Other:
Y/N	Other:
Y/N	Other:

**15/ Have you developed any software for analysis or data extracting?
Please describe the capabilities and outputs of these programs.**

Through its Researchers Service, the BALSAC Project provides help to researchers seeking to use the register. The Service personnel offer all professional and technical expertise as well as resources accumulated over the years in the BALSAC Project (programming consultation, demographic analysis, history, geography, research assistance, population genetics, epidemiology, etc.). The Service intervenes in different ways:

- Technical assistance at various steps of preparation and configuration of research queries so that they are compatible with the structure of the register as well as with the set of privacy guidelines pertaining to the BALSAC Register;
- Preparation of estimations and formatting of the research queries;
- Technical training for accessing data fields in the register;
- Creation of data subsets according to the nature and the goals of the research;
- Reconstitution of genealogies relating to individuals (cases and controls);
- Production of tables and graphs, preliminary analysis (genetic contribution of ancestors, inbreeding and kinship coefficients, etc);
- Loan of data, software, tools, methods developed by the BALSAC Project;
- Etc.

In addition, the BALSAC Project works in collaboration with Helene Vezina, Marc Tremblay and Louis Houde's research group (GRIG) who are developing S-PLUS programs for demogenetic studies based on BALSAC database and its genealogical module RETRO (see <http://www.uqac.ca/grig/>).