

# *PROJET BALSAC*

**Volet : Infrastructures**

**Structure de la base de données BALSAC  
et contenu des principales tables**

/

**Structure of the BALSAC database  
and content of the main tables**

par

JOMPHE Michèle

septembre 2012

**(Document no : I-C-245)**

(adresse postale)

## **PROJET BALSAC**

Université du Québec à Chicoutimi  
555 boulevard de l'Université  
Chicoutimi (Québec)  
G7H 2B1

Téléphone: (418) 545-5517  
Télécopieur: (418) 545-5518  
courriel : [balsac@uqac.ca](mailto:balsac@uqac.ca)

**(English follows at page 7)**

## **Description sommaire de la structure de la base de données relationnelle BALSAC**

La construction du fichier BALSAC se fait en deux étapes principales qui s'opèrent, dans la mesure du possible, selon une démarche chronologique :

1. La saisie informatique des informations contenues dans les actes de l'état civil, suivie d'importantes mesures de validation et de correction.
2. Le jumelage des données : opération destinée à lier et traiter les renseignements relatifs à un même couple de manière à reconstituer les familles. Fondé presque exclusivement sur la comparaison des données nominatives, le jumelage suit une approche progressive au sens où on va des liens les plus faciles (jumelage entièrement automatisé effectué dès l'intégration des informations via le module de saisie) aux liens les plus difficiles (jumelage assisté par ordinateur).

La structure du fichier prend la forme d'une base de données relationnelle, avec une numérotation universelle des individus et des familles, constituée de trois tables principales l'une donnant accès aux individus, la deuxième aux couples, la troisième aux événements (par exemple : naissances, mariages, décès, etc.) tel qu'illustré dans la FIGURE 1.

Une même personne conserve le même numéro à travers chacune des tables.

Dans la table INDIVIDU, chaque personne est reliée à ses parents via les numéros d'individu (IND, PERE, MERE). Dans la table COUPLE, les époux sont également reliés via leurs numéros d'individus (HOMME, FEMME).

Au final, BALSAC est une base de données longitudinale qui permet la reconstitution automatique des histoires familiales, des liens généalogiques et de la structure de la parenté de l'ensemble de la population du Québec sur plus de quatre siècles. Il couvre toutes les régions du territoire québécoise depuis les débuts du peuplement européen, au 17<sup>e</sup> siècle, jusqu'à la période contemporaine.

**Structure simplifiée**  
**de la base de données relationnelle BALSAC**  
**(Trois tables principales)**

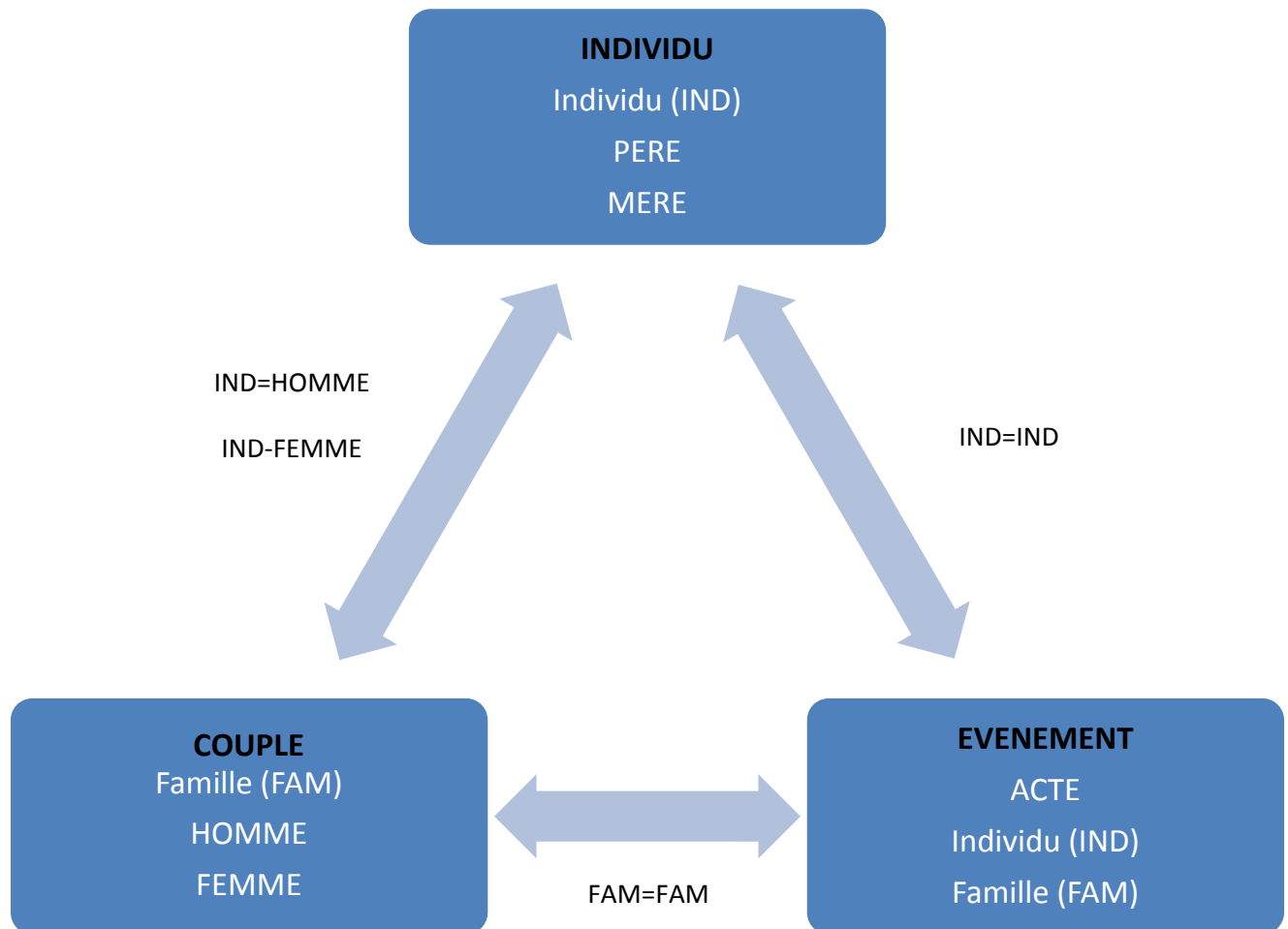


FIGURE 1

## Description des trois tables principales

**TABLE « INDIVIDU »**

<b>CHAMP</b>	<b>FORMAT</b>	<b>SIGNIFICATION</b>	<b>TABLE CORRESPONDANTE</b>
ind	I4	Numéro d'individu	
pere	I4	Numéro du père	
mere	I4	Numéro de la mère	
codenom	I4	Code du nom modal de ind	NOMREEL
codepre	I4	Code du prénom modal de ind	PREREEL
sexe	I1	sexe de ind	SEXE_C
datn	DATE	Date de naissance de ind	
mdatn	I1	Drapeau de précision de daten	MDATE_C
datd	DATE	Date de décès de ind	
mdatd	I1	Drapeau de précision de dated	MDATE_C
aged	I2	Age déclaré au décès de ind	AGED_C
statn	I1	Statut à la naissance de ind	STATN_C
datent	DATE	Date d'entrée en observation de ind	
mdatent	I1	Drapeau de précision de datent	MDATE_C
modent	I1	Mode d'entrée en observation (rôle de ind)	ROLE_C
datsor	DATE	Date de sortie d'observation de ind	
mdatsor	I1	Drapeau de précision de datsor	MDATE_C
modsor	I1	Mode de sortie d'observation (rôle de ind)	ROLE_C
val	I4	Marqueur de l'état de validation	VAL_C

TABLE « EVENEMENT »

CHAMP	FORMAT	SIGNIFICATION	TABLE CORRESPONDANTE
acte	I4	Numéro d'acte	
actem	I1	Numéro de mention	
pos	I1	Position de la personne dans la mention	
fam	I4	Numéro de famille	
ind	I4	Numéro d'individu	
sup	I1	Marqueur de mention superflue	
typact	I1	Type d'acte (baptême, mariage, sépulture, remariage, etc	TYPACT_C
bans	I1	Nombre de bans publiés pour ce mariage	
datenr	D	Date d'enregistrement	
mdatenr	I1	Drapeau de précision de datenr	MDATE_C
datev	D	Date d'événement	
mdatev	I1	Drapeau de précision de datev	MDATE_C
lieuenr	I2	Lieu d'enregistrement	RESID_QC
sexsuj	I1	Sexe du sujet de l'acte	SEXE_C
sexe	I1	Sexe de la personne	SEXE_C
rolind	I1	Rôle de la personne dans l'acte	ROLE_C
codenom	I4	Code du nom de la personne	NOMREEL
codenomv	I2	Variante du nom de la personne	NOMREEL
nom	T64	Nom (en clair) de la personne	NOMREEL
codepre	I4	Code du prénom de la personne	PREREEL
codeprev	I2	Variante du prénom de la personne	PREREEL
prenom	T64	Prénom (en clair) de la personne	PREREEL
resid	I4	Résidence de la personne	RESID_QC
residv	I2	Variante de la résidence de la personne	RESID_QC
prof	I4	Code de profession de la personne	PROF_QC
profv	I1	Variante de la profession déclarée	PROF_QC
honor	I2	Titre honorifique	HONOR_C
ethnie	I2	Code d'ethnie de la personne	ETHNIE_C
origine	I4	Code d'origine de la personne	RESID_QC
originev	I2	Variante du code d'origine de la personne	RESID_QC
em	I1	État matrimonial	EM_C
sign	I1	Présence de la signature	SIGN_C
pres	I1	Présence de la personne	PRES_C
age	I2	Âge déclaré	AGED_C
statn	I1	Statut à la naissance	STATN_C
cons	C4	Consanguinité entre les époux déclarée au mariage	CONS_C
ref	C10	Numéro séquentiel de l'acte dans le registre	
contrat	I1	Marqueur de l'existence d'un contrat	NOTAIRE
obs	I1	Marqueur de la présence d'une observation	OBSERVATION
src	I1	Source de l'acte	SRC_C
val	I4	Marqueur de l'état validation	VAL_C

**TABLE « COUPLE »**

<b>CHAMP</b>	<b>FORMAT</b>	<b>SIGNIFICATION</b>	<b>TABLE CORRESPONDANTE</b>
fam	I4	Numéro de famille	
homme	I4	Numéro d'individu de l'homme	
femme	I4	Numéro d'individu de la femme	
datm	DATE	Date de mariage du couple	
mdatm	I1	Drapeau de précision de datm	MDATE_C
cons	C4	Consanguinité des époux déclarée au mariage	CONS_C
lieum	I2	Lieu de mariage du couple	RESID_QC
datent	DATE	Date d'entrée en observation du couple	
mdatent	I1	Drapeau de précision de datent	MDATE_C
modent	I1	Mode d'entrée en observation (rôle de l'homme)	ROLE_C
datsor	DATE	Date de sortie d'observation du couple	
mdatsor	I1	Drapeau de précision de datsor	MDATE_C
modsor	I1	Mode de sortie d'observation (rôle de l'homme)	ROLE_C
emh	I1	État matrimonial de l'homme	EM_C
emf	I1	État matrimonial de la femme	EM_C
val	I4	Marqueur de l'état de validation du jumelage	VAL_C
jum_statut	I1	Code d'état des opérations de jumelage	JUM_STATUT_C
jum_note	I1	Indication de notes d'observation concernant le jumelage	
jum_statuth	I1	Statut de l'homme déterminé au jumelage	
jum_statutf	I1	Statut de la femme déterminé au jumelage	
jum_origh	I4	Lieu de mariage ou d'origine de l'homme déterminé au jumelage	RESID_QC
jum_origf	I4	Lieu de mariage ou d'origine de la femme déterminé au jumelage	RESID_QC

## Brief description of the structure of the BALSAC relational database

The BALSAC database is developed in two phases, completed in chronological order whenever possible:

1. Data entry: civil registration data is entered, validated and corrected when necessary.
2. Data linkage: all information related to couples is linked and processed to reconstruct families. Based almost exclusively on a comparison of nominative data, linkage follows a progressive approach starting with the easiest step (fully automated linkage process done as soon as the information is transferred with the Data Entry function) to the most difficult (computer-assisted linkage).

BALSAC is a relational database with a universal identification number for individual and families, is comprised of three main tables. One relates to individuals, the second to couples, and the third to events (i.e. births, marriages and deaths), as shown in FIGURE 1.

All individuals keep the same ID number through each table. In the table INDIVIDU, each individual is connected to the parents via ID numbers (IND, FATHER, and MOTHER). In the table COUPLE, spouses are also connected through their ID numbers (MAN, WOMAN).

In the end, BALSAC is a longitudinal database that allows automatic reconstruction of family histories, genealogical and kinship structure of the entire Quebec population over four centuries. It covers all regions of the Quebec territory since the beginning of European settlement in the 17<sup>th</sup> century to the contemporary period.

## Simplified structure of the BALSAC relational database

(Three main tables)

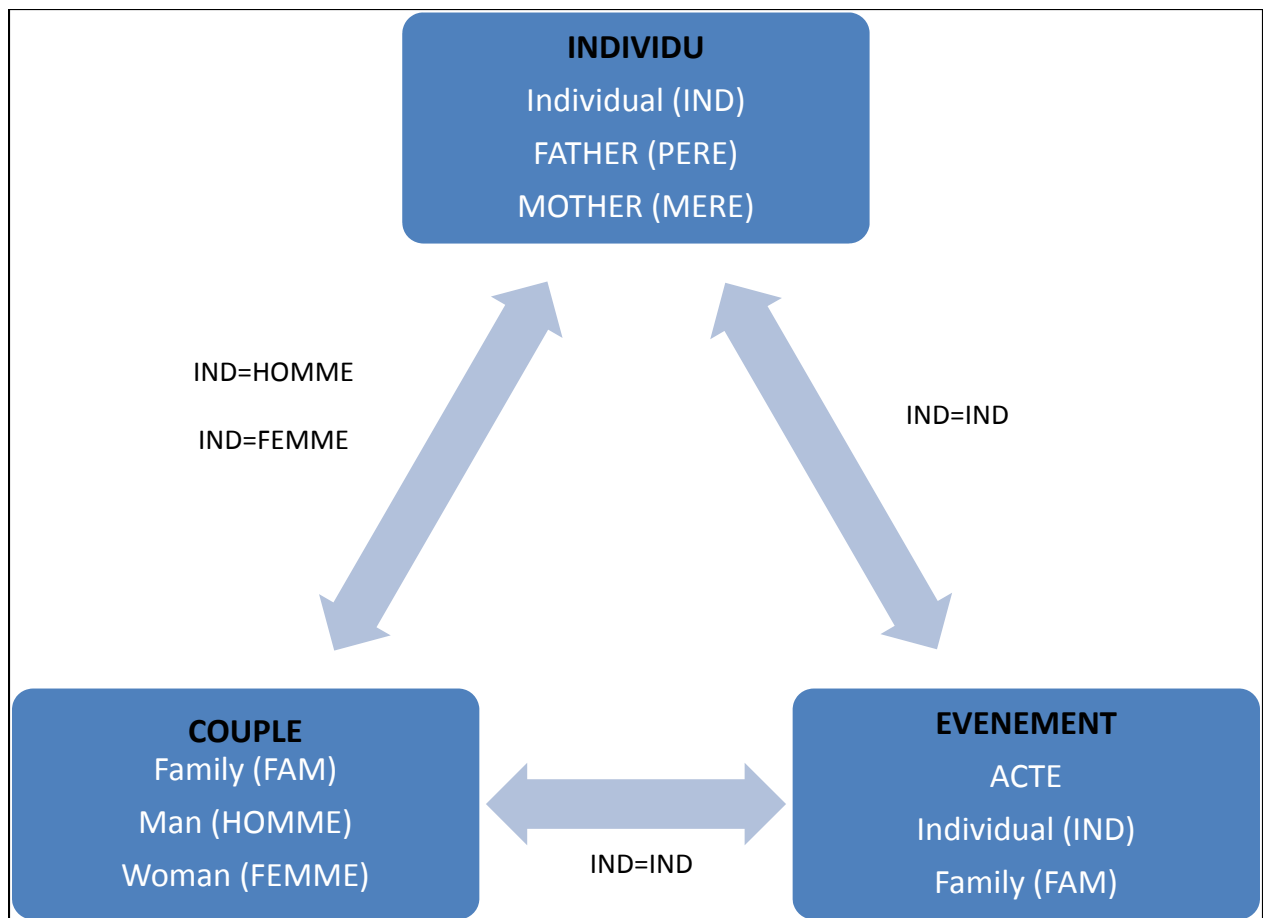


FIGURE 1



## Description of the three main tables

### « INDIVIDU » TABLE

Field	Format	Description
ind	I4	individual identification number
pere	I4	father's identification number
mere	I4	mother's identification number
codenom	I4	surname code
codepre	I4	given name code
sexe	I1	sex
datn	DATE	date of birth
mdatn	I1	flag for datn
datd	DATE	date of death
mdatd	I1	flag for datd
aged	I2	declared age at death
statn	I1	status at birth (for example: illegitimate)
datent	DATE	date of first mention in a vital data record
mdatent	I1	flag for datent
modent	I1	role at first mention
datsor	DATE	date of last mention in a vital data record
mdatsor	I1	flag for datsor
modsor	I1	role at last mention
val	I4	flag for validation status

« EVENEMENT » TABLE

	Format	Description
acte	I4	certificate number
actem	I1	mention number
pos	I1	individual position in the certificate
fam	I4	family number
ind	I4	individual number
sup	I1	flag for redundant mention
typact	I1	type of certificate
bans	I1	number of banns published for a marriage
datenr	DATE	date of registration
mdatenr	I1	flag for datenr
datev	DATE	date of event
mdatev	I1	flag for datenv
lieuenr	I2	place of registration
sexsuj	I1	sex of the subject
sexe	I1	sex of the individual
rolind	I1	role of the individual
codenom	I4	surname code
codenomv	I2	surname variant
nom	T64	surname
codepre	I4	first name code
codeprev	I2	first name variant
prenom	T64	first name
resid	I4	place of residence
residv	I2	place of residence variant
prof	I4	profession
profv	I1	profession variant
honor	I2	honorary title
ethnie	I2	ethnic group
origine	I4	place of origin
originev	I2	place of origin variant
em	I1	matrimonial status
sign	I1	flag for signature
pres	I1	flag for presence
age	I2	age
statn	I1	status at birth
cons	C4	declared consanguinity of spouses at marriage
ref	C10	certificate sequential number in the register
contrat	I1	marriage settlement by a notary
obs	I1	flag for notes and commentaries
src	I1	source of the certificate
val	I4	validation status

« COUPLE » TABLE

Field	Format	Description
fam	I4	family number
homme	I4	husband's identification number
femme	I4	wife's identification number
datm	DATE	date of marriage
mdatm	I1	flag for datm
cons	C4	declared consanguinity of spouses at marriage
lieum	I2	place of marriage
datent	DATE	date of first mention in a vital data record
mdatent	I1	flag for datent
modent	I1	role of husband at first mention in a vital data record
datsor	DATE	date of last mention in a vital data record
mdatsor	I1	flag for datsor
modsor	I1	role of husband at last appearance
emh	I1	husband's matrimonial status at marriage
emf	I1	wife's matrimonial status at marriage
val	I4	validation status
jum_statut	I1	linkage status
jum_note	I1	flag for notes and commentaries pertaining to linkage
jum_statuth	I1	linkage status for the husband
jum_statutf	I1	linkage status for the wife
jum_origh	I4	husband's place of origin
jum_origf	I4	wife's place of origin