Research data access services (RDA)

Startup kit

Checklist table 1
Match file

For the purpose of matching data and producing your project's research file, you must transmit data to the ISQ so that they can be matched to the data extracted by the ISQ.

This checklist table sets out the checks that the holder must complete before transmitting data file(s) to the ISQ.

Important

- **Two separate files** must be supplied to the ISQ for the production of the research file.
  - A match file (identification data);
  - A file containing research variables (without identification data).
- These files must be **transmitted separately** using a secure exchange method (two separate sends).
- Files must not **under any circumstances** be sent via email.
- **You must complete this table and send it with** your match file.

It is the researcher's responsibility to ensure the quality and consistency of information in the files transmitted to the ISQ. The ISQ will not perform this validation.

INSTITUT
DE LA
STATISTIQUE
DU QUÉBEC

Québec

Match file (with identification variables)

Description: The match file(s) contain variables allowing an individual or organization to be identified directly (e.g. last name, first name, address, health insurance number, etc.) and any other information necessary for matching (including the index date).

Checks to be completed by the researcher

	Yes	No	Comments
Anonymous numbers and individuals			
Include only information of individuals who match the defined selection criteria.	◯	◯	
The number of records in the cohort transmitted to the ISQ is , and corresponds to the estimated size of the cohort for cost evaluation purposes.	◯	◯	
The identification file(s) must contain an anonymous number uniquely representing individuals (the number must not be linked to a person's identity).	◯	◯	
The anonymous number must make it possible to locate the same individual present in the researcher's various files.	◯	◯	
The anonymous number must correspond to a single individual in your files.	◯	◯	
Every individual must be represented by a single number.	◯	◯	
If the file(s) contain more than one line for the same individual (e.g. several addresses), the individual's unique number must be repeated for each line.	◯	◯	
If an individual is present more than once in the file(s), please notify us.	◯	◯	
Index date			
If an index date must be taken into account, it must be clearly defined in the file(s) transmitted. *E.g. index date = event 1*	◯	◯	
If there are several index dates, each must refer to a specific event. *E.g. index date 1 → event 1; index date 2 → event 2; etc.*	◯	◯	

Continued on page 3

	Yes	No	Comments
Variables			
File(s) must contain one item of information per column. *E.g. Last name = column 1; First name = column 2, etc.*	◯	◯	
The length of the variable name must not exceed 30 characters.	◯	◯	
The variable must contain nº special characters and nº spaces.	◯	◯	
A description (label) of the name of every variable must be supplied.	◯	◯	
For any one variable, the format must be standardized in the various files to be received (e.g. date).	◯	◯	
All variables necessary for matching, as agreed with your analyst, are supplied.	◯	◯	
File format			
The file(s) must be in one of the following formats: SAS, CSV, Excel. Files in text format (.txt) are acceptable, but not preferred.	◯	◯	
File(s) must be compressed in a password-protected .zip file (the files themselves are not password-protected).	◯	◯	
The filename(s) must contain the lead researcher's last name and must refer either to the file number assigned by the ISQ (e.g. 40001A01) or else to the CAI number assigned to the project, and must contain the word "Appariement". *E.g. 1 (ISQ): "TREMBLAY_40001A01_Appariement"* *E.g. 2 (CAI): "TREMBLAY_#CAI_Appariement"*	◯	◯	
If the size of the compressed file(s) exceeds 35 MB, please notify us.	◯	◯	