



Statistics
Canada Statistique
Canada

Quality Assessment of Vital Statistics

www.statcan.gc.ca



Telling Canada's
story in numbers

Valerie Gaston
Health Statistics Division
Statistics Canada

February 24, 2017
United Nations Expert Group Meeting

Canada 

Introduction of civil registrations and vital statistics in Canada

- 13 provincial and territorial vital statistics registrars.
- Each operates under its own Vital Statistics Act.
- Collection and dissemination of national vital statistics are the responsibility of Statistics Canada (StatCan) which operates under the Statistics Act.
- Use of National Routing System (NRS) to transmit data from provinces and territories to StatCan

Assessing quality - coverage

- NRS is the first step where messages are validated against the schema.
- Reconciliation of counts between electronic data, digitised images of registrations, automated mortality classification software, highest registration number reported by jurisdiction.
- Elimination of duplicates within province or territory, between provinces and territories, over two years.

Assessing quality - editing

- The microdata are run through a series of validation edits.
 - Automated corrections or data conversions have been programmed.
 - Historically, records in error have been corrected or verified by manually.
- Flags are added to each record during the validation phase.
 - Edits have severity levels which will inform next steps.

Assessing quality – review of error correction process

- In order to increase efficiency in this phase of processing.
 - Systematic errors were communicated back to the jurisdiction in order for them to adjust their programs.
 - Certain corrections or conversions programmed based on information in other data elements.
 - Blank age of mother but the DOB field contains data, the age will be derived using the DOB and date of event during conversions instead of imputations. Parity of the mother can be better verified during the editing stage.
 - Province of residence is missing and the postal code is available, the province is derived using first letter of the postal code.

Assessing quality – review of error correction process (cont'd)

- Edits were added to correct logic errors:
 - Verify marital status as “single” for deaths of children under the age of 15.
- Edits were adjusted:
 - Threshold for duration of pregnancy and birth weight ranges were changed to allow for heavier babies not to get flagged for verification.
- Edits were or removed since they added limited value:
 - Unknown sex edit for stillbirths since it is a valid outcome

Assessing quality – review of error correction process (cont'd)

- The final piece of the evaluation will be to measure the value of the corrections on the precision of the estimates. This may lead to further adjustments and streamlining of the error correction process once the cost benefit analysis is complete.

Assessing quality – causes of death

- Separate process and a shared responsibility.
 - The 3 larger jurisdictions maintain their own trained cause of death coding staff and code their own data.
 - StatCan provides cause of death coding for the remaining jurisdictions.
 - StatCan provides training for all cause of death coders and conducts a cause of death review.
 - Invalid cause, rare codes, first time used codes, age and cause correlations, maternal deaths are reviewed.
 - Ensure consistent application of the classification
 - Addresses known problems with the automated mortality classification system.

Assessing quality – imputations

- In order to ensure that data users wishing to use imputed data are all using the same imputed values, imputation has been incorporated into the data processing activities.
 - Births: sex of child, province or territory of residence of the birth mother, and dates of birth of the birth mother and birth father.
 - Deaths: province of residence, sex, age and date of birth of the decedent.
- This will ensure consistency amongst the various subsequent data outputs.

StatCan Guidelines for Validation of Statistical Outputs

- To support the quality of its statistical products and outputs, StatCan has issued a Directive on the Validation of Statistical Outputs which outlines requirements and responsibilities in conducting validation.
- Validation activities can be based on subject matter and processing procedures for the program.

Subject matter validation activities

- Analysis of changes over time for all variables by province/territory from 2000.
- The time series of a particular statistic over a sequence of time points is analysed, looking for a break in the series, or a large jump in the estimates.
- These time series may be affected by changes in collection or processing. As part of the investigative process, it may be necessary to consult with the data providers.

Subject matter validation activities (cont'd)

- Cross tabulations allow for the internal consistency of the data to be verified and to explore underlying characteristics of the estimate.
 - Province or territory of residence by province or territory of occurrence; birth year and death year.
- Coherence analysis based on known current events.

Subject matter validation activities (cont'd)

- Confrontation with other sources. Retuning reports back to the registrar to confirm certain numbers
 - Month of death and sex of the decedent by province or territory of the death; age group and sex of decedent by province or territory of death; birth or stillbirth by month and sex by province and territory of stillbirth; age group of mother and sex of child by province or territory

Subject matter validation activities (cont'd)

- Consultations with stakeholders within StatCan. Attempt to coordinate releases with those of the Demography and Health Indicators to provide users with a complete range of data products on a similar topic. Analysts from each section will review each other's analytical products to ensure estimates are consistent and analysis is coherent.
- Review of The Daily by senior management.

Production validation activities

- Thorough review of production process. Any changes are documented in user guide.
- Coherence analysis based on data quality indicators with corresponding target values.
 - Coverage (quantifying the number of late registrations received after the file cut-off date; birth-death linkages for infants under the age of one; confrontation to births from hospital data)
 - Response rates for key variables (postal code of residence (mother or decedent), underlying cause of death, location of death, duration of pregnancy, and birth weight), location of birth.
 - Imputation rate tracked.

Validation report

- Records and communicates the outcome of the validation activities.
- Perceived or actual data anomalies should be noted along with the investigation that followed and the correction strategy or the reason for accepting the discrepancy.
- This report forms the documentation behind the decision of “fitness for use” of the program estimates.

Planning for validation process

- Several factors should be considered.
 - The cost of validation work in relation to the total cost of the program needs to be taken into account.
 - Validation steps should not unduly get in the way of timeliness.
 - Staff should obtain the proper training and tools, as well as have the necessary experience and knowledge.
 - Time must be taken to establish and maintain proper internal and external partnerships.

Planning for validation process (cont'd)

- Deciding on the scope and extent of the validation activities is based on several factors:
 - Use and impact of the estimates, and the risk of error in a program. Since vital statistics data are used by the Demography division to respond to a legislative requirement, an in-depth validation is required.
 - Since the Vital Statistics Section has recently redesigned the data processing activities, a comprehensive validation is required.

Additional validation best practices

- Creation of a Governance body to consult if significant data quality issues are identified.
- Peer-review meetings are recommended. Program findings can be discussed during these meetings between program analysts and subject matter experts outside the project team.
- Post-mortem discussions with the project team are a valuable opportunity to come up with ways to improve the process for the next round.
- Segregation of production and validation teams. Separation of duties allows for an independent assessment of the estimates.