Methodological notes

Stages of data flow and sources of error

The Adult Criminal Court Survey data contained in this report have undergone numerous types of processing since the original data collection took place in the courtroom. Each process may introduce error, or a difference between the original court event and the final tabulated data. The following is a description of the four basic stages of data flow for the ACCS and the types of errors which may be introduced at each stage.

Stage 1. Collection of the original data

The initial stage involves the recording of court proceedings in some form of docket. The types of errors that may be introduced include the following:

1. Response and measurement error: Provinces and territories or court locations within the same province or territory may differ in terms of court operations, recording procedures, and the interpretation of statutes or legal concepts.
2. Non-response error: Data pertaining to some ACCS variables may not be collected if not required for operational or administrative purposes.
3. Data coding errors: There may be clerical errors in the recording of court proceedings.

Stage 2. Provincial/territorial processing of the original data

Data from court dockets are loaded into the provincial and territorial automated operational systems. Data are then loaded into automated information system files within the jurisdiction. The types of errors which may be introduced at this stage include:

1. Coverage error: Data for certain court locations may not be captured for inclusion in the automated files.
2. Non-response errors: Data pertaining to some ACCS variables may not be captured for inclusion in the automated operational systems, though they exist in the court dockets.
3. Data coding and capture errors: There may be clerical errors or delays in the coding of data on the docket or in the transformation of data from the docket into a machine-readable format.
4. Edit and imputation errors: The edit and imputation process used by the provincial/territorial systems may introduce errors.
5. System errors: Processing of operational system files or information system files may introduce errors.

Stage 3. Transcription of information into ACCS format

Interface software selects the applicable records from automated court operational systems or management information systems in the provinces/territories. Records are then transcribed into the ACCS format and sent to the Canadian Centre for Justice Statistics (CCJS) for processing. The types of errors that may be introduced include:

1. System errors: ACCS interface software specifications may be incorrect or misinterpreted or the interface software itself may contain errors.
2. Coverage errors: The ACCS interface software may extract inappropriate records or miss records that should be covered by the survey.

Stage 4. ACCS processing of the transcribed data

Data received from the provinces and territories are edited and, if necessary, data are imputed before being added to the ACCS master file. Automated standard tables are produced from the contents of these files. The types of errors that may be introduced at this stage include:

1. System errors: ACCS processing specifications may be incorrect or misinterpreted. The processing programs themselves may also contain errors.

2. Edit and imputation errors: Although the edit and imputation system is a tool for detecting and correcting errors, it may also be a source of error. The edit specifications may contain errors due to changes in court procedures, or legislation. The edit and imputation rules may be mis-interpreted or they may be too restrictive or not restrictive enough.

ACCS tools for assuring data quality

A number of tools have been developed for the ACCS to minimize or correct the errors identified above.

A. Interface software development methodology

System error may be introduced during the extraction and transcription of provincial or territorial data into ACCS format. The ACCS has attempted to minimize this source of error by implementing a standard interface development methodology which requires a complete testing of the software by both the ACCS and the province/territory prior to implementation. Interface development must proceed through the following stages:

1. The feasibility of developing interface software is investigated. A test environment is established in the automated operational or case management system of the province or territory to assist in the feasibility investigation and in the development of interface software.
2. Provincial/territorial data elements are mapped to the ACCS data requirements and documented in a field interpretation and limitations(FID) template.
3. The interface software is run in the test environment using test data and the results are verified by both the ACCS and the province/territory. Differences between actual and expected results are identified and the required corrections are made.
4. The interface software is run in the production environment and the results are verified by both the province/territory and ACCS. Errors are identified and the required corrections are made.
5. An evaluation exercise is undertaken to assess the extent to which the data extracted by the interface match the data contained in the original provincial/territorial court files and automated operational or case management system files. The review facilitates the investigation of response errors or discrepancies between the published data and the original recording of the court event.

B. Testing of ACCS data processing systems

The ACCS data processing systems may introduce system errors and edit and imputation errors into the data. To minimize the impact of these errors, all systems developed at Statistics Canada, including the ACCS data processing system, are subject to logic testing by the developer, user acceptance testing performed by the CCJS and/or the Methodology Division of Statistics Canada, and volume testing performed by the system developer. It should be noted, however, that these tests only approximate the infinite number of situations encountered in the real world. Periodic revision will also be required to reflect changes in court procedures and law. For these reasons, the ACCS data processing system requires continuous review and modification.

C. ACCS edit and imputation system

The main tool developed for the ACCS to identify and correct errors and to guide interface software development is an edit and imputation system. Erroneous data processed by the edit system may result in one of the following outcomes:

1. Data are rejected: There are three types of errors that cause data to be rejected: missing or incorrect information in the key record identification fields (Province/Territory, Court Location, Information Number, Accused Identifier, Charge Sequence Number, and Court Appearance Date); charges not yet disposed; and duplicate appearance records. In most cases, records containing reject errors require respondent input to determine the appropriate action to be taken.

2. Data are corrected: A pre-determined correction is made based on the specific combination of data element values. This correction method applies when a value associated with a data element is incorrectly coded or a combination of data elements are logically inconsistent within an appearance record or charge.
3. A warning is issued: This occurs when there are inconsistencies in the combination of data elements or there is incorrect coding.

Data Assessment Procedures

Data assessment activities are undertaken by both the provinces/territories and the ACCS to monitor data quality and to provide direction for any modifications where data quality problems are identified. The data are subjected to year-to-year comparisons. Comparisons are made between the survey data and figures contained in the provincial/territorial reports of court operational or case management systems. The results of the data assessment process permit the identification of most types of errors occurring at all stages of data flow.