Data Quality in the Therapeutic Abortion Survey

Background

The Therapeutic Abortion Survey was originally designed to measure all legally induced abortions performed in Canada (and to the extent possible, abortions performed in the US on Canadian residents prior to 2004). Statistics Canada was responsible for the survey up to the 1994 data year. As of the 1995 data year, the Canadian Institute for Health Information (CIHI) assumed data collection, compilation and processing responsibilities. Statistics Canada remains involved in the approval of the final annual file and plays a major role in the dissemination of data from the survey.

Data sources

There are several data sources and formats used in the creation of the Therapeutic Abortion Survey database. CIHI obtains data from provincial/territorial departments of health, or directly from hospitals and clinics.

Depending upon the source, the format can also vary from a single sheet of paper with aggregate counts to detailed electronic records submitted through the Discharge Abstract Database (DAD).

The reconciliation of various data sources and formats has the potential to affect accuracy. For example, 1) lack of standardization increases the potential for variations in definitions and concepts, such as the calculation of gestation period; 2) an increased risk of data entry errors when processing paper reports; 3) loss of detail when mapping values used in one collection system to that used in the Therapeutic Abortion Survey. A continuous high level of data quality surveillance is necessary.

As of the 1999 data year, the DAD is being used as the data source for hospital abortions for those provinces that submit to the DAD. This move has resulted in increased standardization and quality control.

From Newfoundland and Labrador, data for hospital induced abortions are submitted electronically by hospitals and contain all data elements. Data for clinic induced abortions are aggregate counts submitted by the clinics on paper forms.
Prince Edward Island has not reported to the Therapeutic Abortion Survey since 1983.
From Nova Scotia, data for hospital induced abortions are submitted electronically by hospitals and contain all data elements. As of 2004, there are no longer any abortion clinics operating in Nova Scotia.

From New Brunswick, data for hospital induced abortions are submitted electronically by hospitals and contain all data elements. Data for clinic induced abortions are aggregate counts submitted by the clinics on paper forms.

From Quebec, aggregate counts of hospital and clinic induced abortions are submitted electronically by the Ministry of Health.

From Ontario, data for hospital induced abortions are submitted electronically by hospitals and contain all data elements. Data for clinic induced abortions are aggregate counts submitted electronically by the Ministry of Health.

From Manitoba, data for hospital induced abortions are submitted electronically by the Ministry of Health and hospitals and contain all data elements. In 2004 and 2005, information on abortions performed in clinics in Manitoba was not submitted to the Therapeutic Abortion Survey.

From Saskatchewan, data for hospital induced abortions are submitted electronically by hospitals and contain all data elements.

From Alberta, data for hospital induced abortions are submitted electronically and contain all data elements. Data for clinic induced abortions are submitted electronically by the Ministry of Health and contain all data elements.

From British Columbia, data for hospital induced abortions are submitted electronically by hospitals and contain all data elements. Aggregate counts of clinic induced abortions are submitted on paper forms by the clinics.

From the Yukon Territory, data for hospital induced abortions are submitted electronically by hospitals and contain all data elements.

From the Northwest Territories, data for hospital induced abortions are submitted electronically by hospitals and contain all data elements.

From Nunavut, data for hospital induced abortions are submitted electronically by hospitals and contain all data elements.

Coverage

CIHI estimates that the Therapeutic Abortion Survey database represents approximately 90% of all abortions performed in Canada involving Canadian residents.[1] Coverage improved even more in the 2001 data year due to the submission of counts from a facility in British Columbia that had previously not participated in the survey. Although 90% of abortions are accounted for in the database, many of these abortions are reported only as aggregate counts. See the next section on Unit and Item Response.

At the national level, the percentage of abortions performed on non-residents of Canada that are included in the database is not known. The percentage of abortions performed in the United States on Canadian residents that are represented in the survey's database is also not known.

[1] Canadian Institute for Health Information. Privacy impact assessment of the Therapeutic Abortions Database. June 2003. Page 4.

Limitations to Coverage

1. Medical (pharmaceutical) abortions: With the recent introduction of medical abortions, it is becoming increasingly difficult to ensure the collection of data on all induced abortions. For example, it is not known if medical abortions are being initiated in physician offices, in addition to the traditional locations of hospitals and clinics. The Therapeutic Abortion Survey currently collects information on abortions performed in hospitals and clinics. Furthermore, it is not possible to track a patient once they have been discharged. As such, if a patient is released but later requires another procedure as a result of complications, for example an incomplete abortion, it is possible that she would be counted twice in the database. This issue is particularly relevant with the growing popularity of medical abortions in Canada, which have a higher rate of incomplete abortions than do surgical abortions. Is a woman seeks treatment in a clinic and then requires further intervention elsewhere, that case will be counted twice in the database.

2. Non-residents: For Quebec, prior to 2004, all aggregate counts of abortions received from the Quebec ministry of health were assumed to be for insured residents of that province. Detailed records were received for some hospital abortions and in such cases the residency was indicated. However, as of 2004, only aggregate counts for both hospital and clinic abortions are submitted by the Quebec ministry of health and all counts are assumed to pertain only to insured residents of Quebec. For Ontario, as of the 1999 data year, clinic abortion statistics include insured residents of Ontario only. As well, there is a question of whether the counts of insured abortions have a cap. [See next section on Unit and Item non-response for more details on the current situation in Ontario.] In 2001, private clinics in British Columbia stopped reporting the residency of patients, thus eliminating information on non-residents receiving a clinic abortion in that province. In 2002 and 2003, however, British Columbia also provided broad residency information: Canadian or International. All counts indicated as Canadian were assigned a residence of British Columbia; the international counts were assigned to the non-resident of Canada category. As of 2004, hospital information for British Columbia was downloaded through the DAD which provides residency information. British Columbia clinics reported either Canadian or non-Canadian residency, or did not report a place of residence at all. All counts that indicated Canadian residency or where no residency was reported were imputed to residents of British Columbia.

3. Survey frame: Frame validation and maintenance of the survey is a complex procedure because many data sources are involved. CIHI maintains and validates as much as possible the list of facilities it receives from its data sources (e.g., DAD, ministries of health). Detecting facility non-response is a continuing major challenge due to the many changes made to facility numbers resulting from mergers and closures. Facility-specific frames are not available for data sources that submit counts.

4. United States reports: On average, CIHI received reports from 13 American states, mainly those situated along the Canada-US border. For example, for data year 2003, CIHI received reports from Hawaii, Idaho, Maine, Michigan, Minnesota, Montana, New Mexico, New York state, North Dakota, South Dakota, Oregon, Vermont and Washington state. No reports are received from states such as Massachusetts, Connecticut, Ohio, Pennsylvania, Florida and California. As of 2004, CIHI no longer obtains reports on Canadian women obtaining abortions in the United States.

5. Prince Edward Island has not reported to the Therapeutic Abortion Survey since 1983, mainly because non-emergency abortions are not allowed in Prince Edward Island. Cases of the occasional emergency abortion are not reported.

6. Survey not treated as mandatory: coverage of abortions performed in Canada was considered to be 100% prior to 1988. In January 1988, the Supreme Court of Canada struck down the 1969 abortion law, and some hospital and provincial ministry respondents interpreted this action as the basis for no longer having to report to the Therapeutic Abortion Survey. The law had included a provision enabling provincial ministries of health to obtain abortion data from hospitals. At the federal level, however, Statistics Canada surveys (including the Therapeutic Abortion Survey) are mandatory unless otherwise specified, but Statistics Canada chose to treat the Therapeutic Abortion Survey as 'voluntary' but encouraged respondents to continue to supply data for health-related purposes.

7. Nunavut reported data for only the first three months of 2002 and the last nine months of 2003. As a result, statistics for those years exclude Nunavut.

8. In 2004 and 2005, information on abortions performed in clinics in Manitoba was not submitted to the Therapeutic Abortion Survey.

9. For the 2003 data year, a facility in British Columbia did not report to the Therapeutic Abortion Survey. Counts were estimated for this facility based on the number of abortions it reported in previous years. The same methodology was applied to a different facility in British Columbia that did not report in 2005.

Unit and Item Non-response

In addition to complete non-response, another type of non-response is the provision of aggregate counts of abortions rather than detailed records. The submission of aggregate counts is the major limitation of the Therapeutic Abortion Survey.

Text table 1
Induced abortions, by type of facility

Reference year and type of facility	Number of abortions	Percentage with detailed records
1998		
Hospital	68,290	69
Clinic	42,230	48
Total	**110,520**	**61**
1999		
Hospital	63,832	71
Clinic	42,030	11
Total	**105,862**	**47**
2000		
Hospital	63,535	72
Clinic	42,134	11
Total	**105,669**	**48**
2001		
Hospital	61,259	71
Clinic	45,239	12
Total	**106,498**	**46**
2002		
Hospital	58,536	72
Clinic	46,999	16
Total	**105,535**	**47**
2003		
Hospital	56,089	72
Clinic	47,530	13
Total	**103,619**	**45**
2004		
Hospital	53,758	69
Clinic	47,005	13
Total	**100,763**	**43**
2005		
Hospital	50,562	68
Clinic	46,692	15
Total	**97,254**	**43**

The number of submissions of aggregate counts instead of individual records is increasing. In reference year 1998, only 61% of total reported abortions had individual records. In 1999 this percentage fell to 47%, largely due to the absence of detailed records for abortions performed in clinics in Ontario. In 2005, 43% of abortions were reported with detailed records.

The Therapeutic Abortion Survey identifies the type of facility in which the abortion was performed (i.e., hospital or clinic). The table shown above indicates that the increased submission of aggregate counts is mainly a problem with clinic abortions. This is definitely a concern, especially since the proportion of clinic abortions to total abortions is rising. In 1996, abortions performed in clinics represented 33% of total abortions and rose to 48% in 2005.

All hospital abortions, with one exception, are reported in individual record format. In 2005, none of the hospital abortions performed in Quebec was reported with detailed records.

On the clinic side, as of 1999, Alberta became the only province that submits detailed records. Previously, Ontario had been the only other province that submitted detailed records for clinic abortions. In 1999, Ontario began to report only aggregate counts of clinic abortions based on the provincial billing information. Prior to 1999, Ontario's data collection methodology provided complete demographic and medical information on women obtaining abortions in clinics in that province. A comparison of data obtained from the old and new sources for the years 1995 to 1998 revealed undercoverage of about 5.5% per year. The undercoverage was mainly attributed to uninsured services, which are excluded from the counts. The new data source is based on claims made by the physicians, which does not include cases where the patient has paid out-of-pocket for the services. It also excludes cases of residents from other provinces obtaining abortions in Ontario.

Item Non-Response

The analytical data elements found on the detailed record are:
- Province of report
- Facility information (clinic or hospital)
- Province of residence
- Age in single years
- First day of last menses or gestation in weeks
- Date of abortion
- Inpatient days of care
- Number of previous deliveries
- Number of spontaneous abortions
- Number of induced abortions
- Initial procedure
- Subsequent procedure
- Type of sterilization
- Complications

As illustrated in the previous table, detailed records were submitted for only 43% of abortions performed in 2005. Detailed records have virtually 100% response at the item level.

In cases where only counts are submitted instead of detailed records, the counts are aggregated by age group and/or surgical procedure. For example, Ontario counts of clinic abortions are aggregated by age-group, and both hospital and clinic abortions performed in Quebec are aggregated by age-group.

Impact of item non-response

a) Area of residence

Until 2004, only aggregate counts were supplied for abortions performed in British Columbia. The two sources of data from that province were the ministry of health and private clinics. The ministry of health supplied counts aggregated by age group and initial procedure. Records were generated from these counts, and a residence of British Columbia was assigned to all records. [In 2002 and 2003, however, British Columbia also provided broad residency information: Canadian or International. All counts indicated as Canadian were assigned a residence of British Columbia; the international counts were assigned to the non-resident of Canada category.] Prior to data year 2001, the private clinics reported the number of non British Columbia residents. In 2001, however, the clinics no longer supply residency information, thus all records were assigned a residence of British Columbia. In 2002 and 2003, however, British Columbia also provided broad residency information: Canadian or International. All counts indicated as Canadian were assigned a residence of British Columbia; the international counts were assigned to the non-resident of Canada category. As of 2004, hospital information for British Columbia was downloaded through the DAD which provides residency information. British Columbia clinics reported either Canadian or non-Canadian residency, or did not report a place of residence at all. All counts that indicated Canadian residency or where no residency was reported were imputed to residents of British Columbia.

b) Age group imputation:

Age group is the only other data element for which a value other than unknown or unspecified is imputed. This information is required for input into the calculation of pregnancy statistics, especially teen pregnancy.

- Aggregate counts by age group are supplied by the ministry of health in Quebec for hospital and clinic abortions. As of 1999 the ministry of health in Ontario supplies aggregate counts by age group for clinic abortions. In British Columbia, aggregate counts were supplied by age group from 1999 to 2003. For clinic counts where no age group breakdown is provided, an age group is imputed using the known age group distribution observed in hospital abortions. Counts from clinics that do not supply age group constitute approximately 6% of total abortions performed in Canada in 2005.

- Age in single years (which is available only on detailed records) is important in the calculation of teenage pregnancy. Respondents who submit aggregate counts

by age group provide counts for only one teen group: 15-19. However, teenage pregnancy statistics are presented in the subgroups 15-17 and 18-19 years of age because the experiences of the 15-17 age-group are known to be different from the 18-19 age-group. Therefore, estimations are done to subdivide the 15-19 year old age group that is reported or imputed in the Therapeutic Abortion Survey database. Estimations are done using the same methodology used to impute age groups in the database; that is, the known age distribution (in single years) for hospital abortions is applied to those abortions that were reported as counts.

- For Prince Edward Island residents obtaining an abortion in another province/territory that reports only counts by residency, the national age distribution is used.

- Counts of abortions performed in the United States on Canadian women are submitted by age group. In 2003, there were 149 reports received from the US. As of 2004, CIHI no longer receives reports on Canadian women obtaining abortions in the United States.

c) Diminishing Core Data Set

Although the survey's database contains some demographic and medical information, it is increasingly difficult to respond to requests since only 43% of total abortions contain detailed information. Requests for information on, for example, the percentage of reported complications by age group of mother, or gestation period of the fetus can only be fulfilled by using detailed records. Therefore, clients are always informed that such statistics are based only on detailed records. This is more of a problem when looking specifically at clinic abortions.

The current challenge is to provide at least a minimal data set at the national level. Beyond this goal, there are many client requests that the survey cannot meet. For example, the survey does not collect data elements such as education or income level of mother, reason for the abortion, and any complications arising after discharge.

Identifying and correcting errors

Detailed records submitted to the survey undergo an edit process. The edit system checks for internal consistencies, compatibilities and completeness of each data item reported. There are 29 edits and 39 cross-edits. The edit system is reviewed and updated periodically. In calendar 2005, the number errors detected at the initial edit process was 1,564 (3.8%) out of 41,425 detailed records. These errors consist of ‘hard’ errors in which the reported values are invalid and ‘soft’ errors in which the reported values should be confirmed.

Until 2004, error reports were generated for the facilities that supplied detailed records. These reports were sent to the applicable ministries of health. If corrections were provided, revisions to the data were made. In the absence of any further clarification,

invalid codes were changed to a default value 'Unknown'. In calendar 2003, there were 361 records (0.8% of total detailed records) that contained invalid data defaulted to ‘Unknown’.

Starting in 2004, the error reports were not sent to the Ministries and /or facilities. Instead, the erroneous data was reformatted to pre-determined default values.

Upon receipt of a clean file and supporting documentation from CIHI, Statistics Canada carries out a series of quality checks that include: 1) reviewing the record layout and data processing reports, 2) producing basic tabulations to ensure that STC and CIHI results are the same, 3) checking for internal consistencies, e.g. running frequencies on certain data elements, and 4) comparing the most recent data year with past data years to detect any unusual or unexpected changes. After Statistics Canada approves the final file, a public release is announced in *The Daily*.

Other Accuracy Issues

1. Internal provincial discrepancies have been detected in cases where there are two data sources for the same abortion event. Prior to 2004, in British Columbia and Quebec, some clinic abortions were reported by the respective ministry of health and by the clinic that performed the abortion. CIHI reconciled these numbers, and if a discrepancy was found, the higher number was used. As of 2004, this is not a problem: hospital abortions for British Columbia are downloaded from the DAD and clinics submit their counts directly to CIHI; and for Quebec, the ministry of health is the single source of aggregate counts for all abortions.

2. The gestational age is about a week later when the gestation age is calculated from date of last menses than when gestational age is reported as the number of weeks of gestation. The problem is with the rounding method used in calculating weeks rounded from the date of last menses. This problem is significant, as gestation period appears to be gradually shifting down as more records (especially those obtained from the download from DAD) have clinical gestation reported rather than date of last menses.)