

A Critical Review of the Data Quality Report for the Canadian Cancer Registry (CCR)

Ms. G Noonan^{1,3,4}, Ms. D Zakaria², Ms. P Murison^{3,4,5}, Mr. T Snodgrass^{3,4,5}, Dr. L Perron^{4,7}, Mr. M Levesque^{4,8}, Ms. H Anderson^{2,3,4}, Ms. MJ King^{3,4,6}



¹CancerCare Manitoba, Winnipeg, MB, Canada. ²Canadian Cancer Registry, Statistics Canada, Ottawa, ON, Canada. ³Data Quality Management Committee, Ottawa, ON, Canada. ⁴Data Quality Indicators Workgroup, Ottawa, ON, Canada. ⁵Alberta Health Services, Calgary, AB, Canada. ⁶CancerCare Ontario, Toronto, ON, Canada. ⁷Institut national de santé publique du Québec, Québec, QC, Canada. ⁸New Brunswick Cancer Network, Fredericton, NB, Canada.



BACKGROUND

- The Canadian Cancer Registry (CCR) is a national dynamic administrative survey established in 1992, which contains person-oriented demographic and clinical information on primary tumours diagnosed in Canadian residents.
- Thirteen provincial/territorial cancer registries (PTCRs) submit data to the CCR.
- The CCR has implemented a variety of measures to optimize data quality including the annual production of a high level data quality indicator report which allows for comparisons of key indicators across PTCRs.

Original Data Quality Report

green = optimal yellow = good performance red = needs improvement

Indicators	Metric	Optimal Value	PTCR#1	PTCR#2	Etc.
1. Completeness of Case Ascertainment	%	≥ 90	91.0	88.0	...
2. Records Rejected by Edit System	%	≤ 1	0.9	1.4	...
3. Microscopically Confirmed Cases	%	≥ 93	96	93	...
4. Death Certificate Only	%	≤ 3	3.8	4.0	...
5. Unknown Primary Site of Cancer	%	< 2.0	2.0
6. Missing Information					
Racial code	%	≤ 1	0.1	0.3	...
Month of diagnosis	%	≤ 1	0.1	0.1	...
Month of birth	%	≤ 1	0.1	0.9	...
Death registration number	%	≤ 10	0.1	7.0	...
7. Reporting of Staging Data using the Collaborative Staging System	Yes/No	Yes	Yes	Yes	...
8. Reporting of Cancer Records using ICD-O-3	Yes/No	Yes	Yes	Yes	...
9. Reporting of Multiple Primaries (CCR rules)	Yes/No	Yes	Yes	Yes	...
10. Data Submission Delay (Timely reporting of data)	Yes/No	Yes	Yes	Yes	...
11. Cancer Incidence Fully Reported	Yes/No	Yes	Yes	Yes	...

Purpose

To critically review all available data quality indicators to ensure indicators on the CCR data quality report:

- are the most relevant;
- represent a cross section of data quality domains;
- are clearly defined and correctly calculated;
- have appropriate optimal targets; and,
- are properly interpreted.

Data Quality Domains

- Relevance** – degree to which data meet the real needs of users.
- Accuracy** – degree to which data correctly describe the characteristic being measured.
- Timeliness** – delay between the end of the reference period to which the data pertain and the date the data become available.
- Accessibility** – ease with which data can be obtained.
- Interpretability** – availability of supplementary information necessary to interpret and utilize data appropriately.
- Coherence/Comparability** – degree to which data can be successfully integrated with other statistical information.



METHODS

- A template was developed to ensure a standardized, non-biased, comprehensive review process, including adequate consultation with stakeholders.
- Representatives from small, medium, and large PTCRs and the CCR considered each indicator's overall score and the availability of alternative indicators when making a recommendation to retain, modify or move the indicator to a PTCR operational level report.

Template Used To Evaluate Indicators

INDICATOR NAME (e.g. COMPLETENESS OF CASE ASCERTAINMENT)	Criteria	Comment	Score
1.	Is the indicator clearly defined with respect to... a) ...how it is calculated; and/or b) ...how it is presented for interpretation (e.g. percentage, proportion)?		/1
2.	Does the indicator have an objective interpretation?		/1
3.	Was the indicator designed to measure/quantify a data quality issue (i.e., related to one of the key components of quality such as completeness, accuracy, timeliness, and comparability)?		/1
4.	Are the cut-points or optimal values of the indicator... a) empirically based; and/or b) established by external recognized standard setters (e.g. NAACCR, SEER, NCI, IMCC)?		/1
5.	Does the indicator capture what it purports to capture?		/1
6.	Is the indicator the best measure to quantify the data quality issue?		/1
7.	Is the indicator logical, relevant or current?		/1
8.	Is the indicator actionable?		/1
	Score out of a possible 10 points.		/10

RESULTS

- Several recommendations for the new data quality report were made including the addition of a new indicator: "unknown laterality".

Data Quality Domain	Current Indicator	Recommendations
Accuracy	1. Completeness of Case Ascertainment	Include with modification.
	2. Records Rejected by Edit System	Include with modification.
	3. Microscopically Confirmed Cases	Include with modification.
	4. Death Certificate Only (DCO)	Include with modification.
	5. Unknown Primary Site of Cancer	Retain.
	6. Missing Information	Move to operational level report.
Comparability	7. Reporting of Staging Data Using the Collaborative Staging System	Include with modification.
	8. Reporting of Cancer Records Using ICD-O-3	Move to operational level report.
Timeliness	9. Reporting of Multiple Primaries (CCR rules)	Retain.
	10. Data Submission Delay	Include with modification.
	11. Cancer Incidence Fully Reported	Move to operational level report.

Some highlights follow...

Completeness of Case Ascertainment

Definition: Percentage of expected invasive primary tumours (including in situ bladder) in a registry's surveillance population that are actually captured by the registry.

Primary Limitations: Complex to calculate and understand; excludes prostate and breast cancer because they do not satisfy underlying assumptions; and, questionable underlying assumptions.

Assumptions: Age-standardized sex and cancer-specific incidence to mortality rate ratios are relatively consistent across regions and stable, i.e., not increasing or decreasing.

Recommendations:

- Adopt NAACCR's optimal values.

	OLD	NEW
Green	≥90.0	≥95.0 (gold)
Yellow	80.0 to 89.9	90.0 to 94.9 (silver)
Red	<80.0	<90.0

- Inform users of limitations.
 - Underlying assumptions of indicator.
 - Exclusion of male prostate cancer and female breast cancer (approx. 27% of cancers).
 - Completeness at an aggregate level does not guarantee completeness for specific cancers.
 - Completeness can exceed 100% because of the underlying assumptions of the indicator.

Microscopically Confirmed Cases

Definition: The percentage of tumours confirmed by microscope.

Primary Limitation: A very high percentage (e.g., 99%+) may indicate over reliance on pathology reports and missing clinically confirmed or death certificate only (DCO) cases.

Recommendations:

- Caution users regarding potential case completeness issues associated with a very high indicator score.
- Modify cut-points as follows:

	OLD	NEW
Green	≥93.0	92.0 to 96.0
Yellow	90.0 to 92.9	90.0 to 91.9 or 96.1 to 98.0
Red	<90.0	<90.0 or >98.0

Death Certificate Only (DCO)

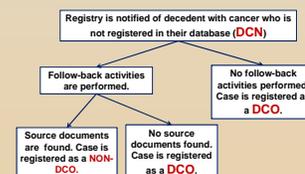
Definition: The percentage of tumours registered by DCO.

Primary Limitations: Although lower scores are better, a value of 0% suggests lack of death clearance activities and thus missed cases. Further, rigorous follow-back activities can drive this indicator towards 0%; the false perception is high case completeness but reality indicates over-reliance on DCOs.

Recommendations:

- Caution users regarding the meaning of 0%.
- Consider feasibility and usefulness of capturing whether a case is death certificate notified (DCN) and monitoring "percentage of DCN tumours."

Death Certificate Only (DCO) vs. Death Certificate Notified (DCN)



Reporting of Staging Data Using the Collaborative Staging System

Definition: Indicates whether a registry submitted at least one tumour with staging data (yes/no format).

Primary Limitation: Does not discriminate between registries with high and low proportions of staged tumours.

Recommendations:

- Redefine as the "percentage of eligible invasive tumours with a valid collaborative stage group."
- Calculate for the top four sites combined and all sites combined.
- Adjust the cut-points based on CCR data and stakeholder goals.

	OLD	NEW
Green	At least one tumour with staging information.	≥90.0%
Yellow	NA	75.0% to 89.9%
Red	No tumours with staging information.	<75.0%

Unknown Laterality

Data Quality Domain: Accuracy.

Definition: The percentage of paired site tumours with an unknown laterality.

	NEW
Green	≤ 5.0%
Yellow	5.1% to 8.0%
Red	> 8.0%

Rationale for Inclusion: Laterality is important for identifying multiple primaries and current indicator scores show room for improvement for the majority of PTCRs.

CONCLUSIONS & NEXT STEPS

- A standardized, comprehensive method of evaluating data quality indicators was developed and successfully implemented.
- The data quality report will be reviewed annually to ensure the most current information regarding key indicators, optimal cut-points, and appropriate interpretation are incorporated.
- Registries have agreed to sharing data quality indicator scores for benchmarking and identifying best practices.

