

BORN Ontario's Data Quality Framework

At BORN Ontario we make *Data Privacy* and *Data Quality* our highest priority. We recognize that the quality of the data directly impacts use of the data. With addition of benchmarking and comparators, the quality of the data across all organizations, practice groups and programs that must be tracked and maintained. Our goal is to ensure that data entered into the BIS is timely, accurate, and comparable as well as usable and relevant for stakeholders.

The BORN Ontario Data Quality Framework (DQF)

The *BORN Data Quality Framework (DQF)* was implemented in 2013. The BORN DQF is based on 5 dimensions of data quality which in turn are divided into 19 elements and sub-elements. Detailed breakdown of the dimensions, elements, sub-elements and definitions are provided in Figures 1 and 2.

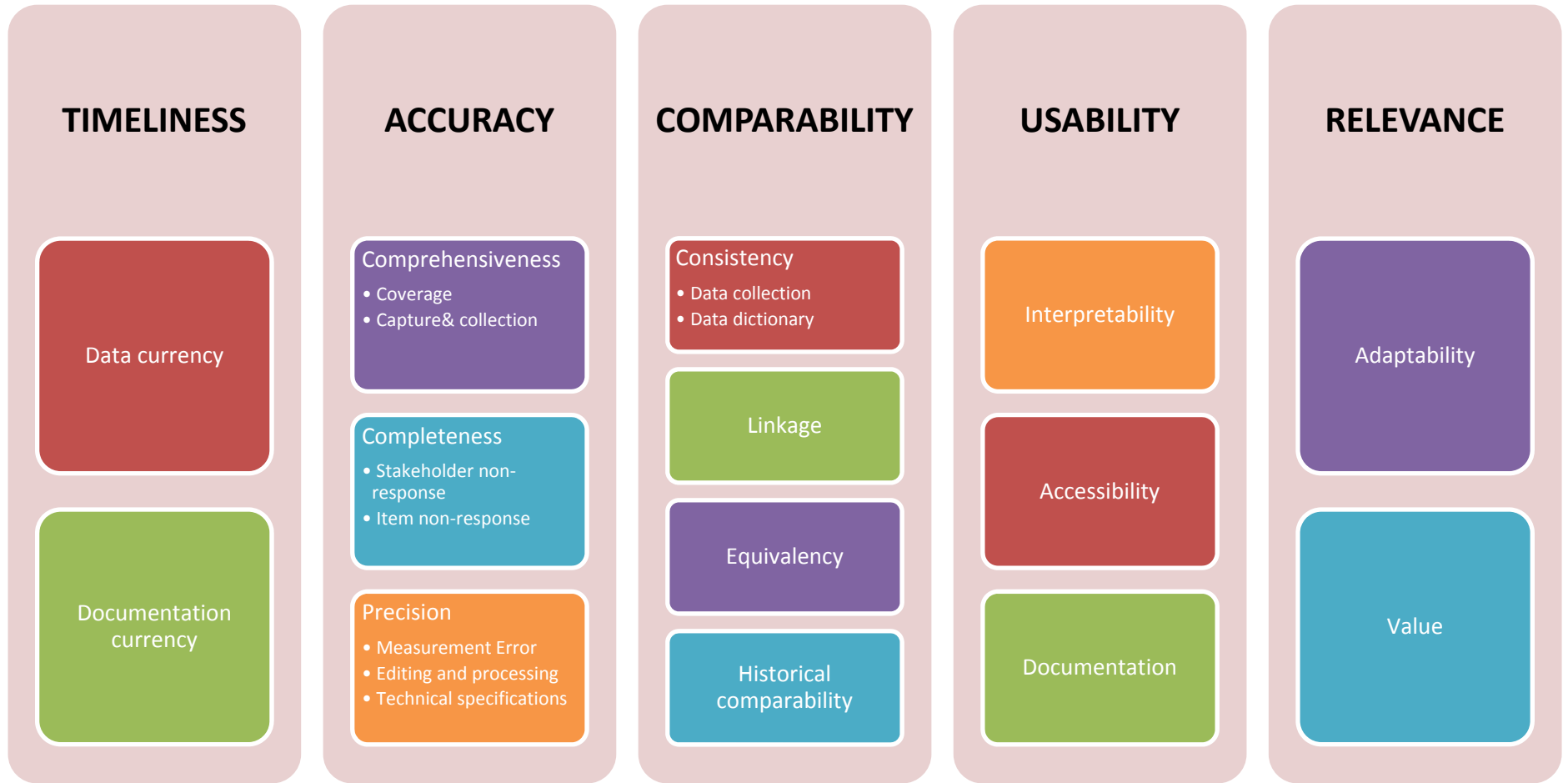
1. **Timeliness** – How current or up to date data is at the time of release and whether the data is available for user needs within a reasonable time period.
2. **Accuracy (validity)** - How well information within or derived from the database reflects the reality it was designed to measure.
3. **Comparability (reliability)** - The extent to which data are consistent over time and entered using standard conventions making it comparable to other databases.
4. **Usability** – The ease with which the data is understood and accessed.
5. **Relevance** – The degree to which the data meets the current and potential future needs of users.

Assessing 'fitness for use' of information in the BIS involves adequately managing each dimension. Failure in one dimension might undermine the usefulness of a final report or data release. The building blocks for this framework require that the data:

- a) Be as current as possible so that decisions are made with recent information,
- b) Be accurate, and entered using consistent conventions to ensure its comparability to like data,
- c) Be accessible and easy to understand, and
- d) Meet the current and potential future needs of the users.

If these dimensions of data quality are achieved, user decision-making will then be based on current, valid, reliable and relevant data.

Figure 1: BORN Data Quality Framework – Dimensions, Elements and Sub-Elements



The BORN Information System brings this framework to bear using a number of features and functions.

Feature	Description	Dimension of Quality
BORN Data Entry Guidelines	Near real-time data entry encouraged	Timeliness (data currency)
BORN Coordinator Support	Monthly acknowledgement reports; Assess and support for data collection response burden	Accuracy (comprehensiveness – capture and collection); Timeliness (data currency);
Robust Reporting Environment	Allows users access to their real-time data (value added feature)	Accuracy (comprehensiveness – capture and collection)
HelpDesk	Prompt response to data collection issues	Accuracy (comprehensiveness – capture and collection)
Drop-Down Lists	Free text is minimized.	Accuracy (precision); Comparability (consistency)
Mandatory Fields & Data Validation Rules	Missing data is minimized	Accuracy (comprehensiveness – capture and collection; precision)
Automated Algorithms	Point of entry validation (e.g. ID is appropriate length and format)	Accuracy (comprehensiveness – capture and collection; precision)
Consistency Edit checks	Point of entry validation (e.g. verifying an intervention that can only be one type of answer)	Accuracy (comprehensiveness – capture and collection; precision)
Dual entry of variables	Point of entry validation (e.g. BW and GA)	Accuracy (comprehensiveness – capture and collection; precision)
Data Capture Quality Control Measures	Point of entry validation (e.g. verification reports, data discrepancy reports, incomplete record reports)	Accuracy (completeness; precision)
Track Missing Data	Stakeholder group and item non-response; Identified in all reports and communications	Accuracy (completeness; comprehensiveness – coverage)
Standard Operating Procedures	To ensure consistency in use of data and analysis methods	Accuracy (precision – processing, editing and estimation)
UAT Testing	To ensure validation of system components (e.g. new encounters and report specifications)	Accuracy (technical specifications)
Data Saved in a Secure Location	Protection of PHI and controls data access	Accuracy (technical specifications)
BORN BIS Data Dictionary	Definitions aligned across stakeholder groups	Comparability (consistency of data); Usability (documentation)
Pre-Population of Data	To improve consistency and reduce duplication and error	Comparability (consistency of data); Accuracy (precision)
Linkage of Cases	Across encounters and to other datasets	Comparability (linkage)
Feedback Mechanisms/Support	BORN Coordinators; Helpdesk	Usability (interpretability)
Data Access Protocols	Key stakeholder; researchers or other users	Usability (accessibility)
BORN Committees	Provide advice about changes for data elements	Relevance (adaptability)

What we are good at:

- *User support* – BORN Coordinators; HELPDESK
- *Guidelines* - developed and available on the BORN website
- *Comprehensive training* – training webinars developed and available on the BORN website for each system component
- *Technical specifications* – developed, documented and used along with UAT testing for new components
- *Robust reporting environment* – available to users with standard reports and enhancements
- *Data Access Protocols* – developed for authorized users (organizations/groups)
- *Dynamic, user-friendly data entry system* - with point of entry validation

Areas that are known to be problematic:

- Tracking data entry and validation times
- Documenting under/over coverage
- Rating and reporting on the magnitude of item non-response (for selected data elements and MND KPIs)
- Monitoring data requests that were not executed due to coverage concerns
- Regularly comparing BIS data with external data source (e.g. CIHI) for live births, stillbirths, NICU admissions
- Identifying core elements for trend analysis (e.g. KPI rates, accuracy and 'missingness')
- Consistently capturing issues about BIS data elements from key stakeholders, users and researchers and processing solutions in an efficient manner

List of improvements being developed and implemented:

- Development of an Annual DQ Report
- Development of a Standard BIS Use Tracking Report
- Plan for systematic external audits (re-abstraction studies) to assess validity of selected data elements
- Development of a user satisfaction survey to evaluate relevance of data elements, system function, access and usability of the data, and user experience
- Assess HELPDESK calls to evaluate support needs and user experience
- Investigate the possibility of incorporating the provider registry (CAPE) into BORN for another level of identifier
- Develop and maintain a BORN document of historical changes - providing an audit trail of BIS development, testing, changes and modifications.
- Develop a process to identify and evaluate new data elements, retiring old elements, revise current elements (including critical appraisal of the literature about potential new data elements and assessment of the feasibility of collecting valid and reliable data)
- Develop a development process to identify clinically meaningful, feasible to measure and actionable data elements for future dashboards
- Develop an online request process with a shopping cart format to facilitate access to data for BORN users
- Develop a document that outlines the steps necessary (technical and communication) when planning changes to deal with an emerging issue that requires a data change

Figure 2: Data Quality Dimensions, Elements and Criteria

Dimensions	Elements	Criteria
TIMELINESS How current or up-to-date the data is at the time of release and whether the data is available for user needs within a reasonable time period	Data currency	The difference between the actual date of data capture and data being available (data entry time and validation time) Data processing activities are reviewed and documented yearly to ensure timeliness
	Documentation currency	Data quality documentation is available at the time of data or report release
ACCURACY (VALIDITY) How well information within or derived from a database reflects the reality it was designed to measure	Comprehensiveness (coverage; capture and collection)	COVERAGE: Under or over coverage = difference between the population of interest or the group of units for which information is wanted (e.g. all hospitals in Ontario, all midwifery practices, all PS labs, all IVF clinics etc.) and the population of reference or the group of units for which the statements are made.
		CAPTURE AND COLLECTION: Standard data-submission procedures exist, are documented, and are followed by data providers to minimize response burden Data-capture quality control measures exist and are implemented by data providers
	Completeness – Missing Data (stakeholder non-response; item non-response)	STAKEHOLDER GROUP NON-RESPONSE: All stakeholders groups/organizations have submitted data
		ITEM NON-RESPONSE: All the submitted records are complete
	Precision (measurement error; processing, editing and estimation; technical specifications)	MEASUREMENT ERROR: Error caused when a data element is coded or answered incorrectly
		PROCESSING, EDITING AND ESTIMATION: All collected data elements are checked for validity and invalid data is flagged. The checks and modifications to the data are logical and consistent
		TECHNICAL SPECIFICATIONS: Technical specifications for the BIS are maintained to allow easy validation of the systems, programs, and applications. Systems are tested when changes are made. Raw data is saved in a secure location

Dimensions	Elements	Criteria
<p>COMPARABILITY (RELIABILITY)</p> <p>The extent to which data are consistent over time and entered using standard conventions making it comparable to other databases.</p>	<p>Consistency (Data collection standards; Data Dictionary Standards)</p>	<p>DATA COLLECTION STANDARDS: Core data elements are collected with the necessary detail required for linking or comparison purposes. For derived data elements, the original data element remains accessible</p>
		<p>DATA DICTIONARY STANDARDS: A data dictionary exists and all data elements are evaluated to determine their inclusion within the BORN Data Dictionary</p>
	<p>Linkage</p>	<p>Cases entered into the BIS are identifiable by either postal code (all six digits) or the relevant Standard Geographic Classification (SGC) (e.g. census tracts).</p> <p>Identifiers are used to differentiate facilities or organizations uniquely for historical linkage. This criteria examines whether linkage is possible not whether linkage is done</p>
	<p>Equivalency</p>	<p>Crosswalk and conversion tables exist documenting issues related to data conversion (e.g. cross walk tables for many to one or one to many relationships; conversion tables for pounds to kilograms or height/weight to BMI)</p>
	<p>Historical comparability</p>	<p>Documentation on historical changes to the data holding exists and is easily accessible</p>
<p>USABILITY</p> <p>The ease with which the data is understood and accessed.</p>	<p>Accessibility</p>	<p>A final analysis data set is made available per fiscal year. Standard tables and analyses using standard format and content are produced</p>
	<p>Documentation</p>	<p>Data quality documentation exists to give internal and external users sufficient information to decide if the quality of the data is appropriate for their intended use</p>
	<p>Interpretability</p>	<p>A mechanism is in place whereby key users can provide feedback to, and receive notice from BORN. Revision guidelines are available and applied to BIS update</p>
<p>RELEVANCE</p> <p>The degree to which the data meets the current and future needs of user</p>	<p>Adaptability</p>	<p>Mechanisms are in place to keep stakeholders stay informed of developments, funnel suggestions from end users and BORN committees and critically apprise the literature to assess the validity of current data collected. Future BIS modifications can be easily made</p>
	<p>Value</p>	<p>The mandate of the data holding fills a health information gap</p> <p>Level of data usage is monitored and user satisfaction is periodically assessed (e.g. internal analysts, external stakeholders and users, client support hotline, FAQs)).</p>