Assessing the Quality of Data Abstracted From Medical Records

High-quality data are essential for effective use of clinical registries.

How is the quality of abstracted data currently measured?

27 unique terms have been used to describe the quality of data abstracted from electronic health records (EHRs),[1] often interchangeably[2] and with conflicting meanings.[1] It's not surprising that the discussion of data quality for clinical registries is often confused.

Given the current lack of consensus on how data quality assessments should be conducted,[3] one way to clarify the situation is to ensure the consistent use of definitions.

Definitions have been established for a core set of quality dimensions for health data.

We recommend a core set of complementary measurements for a comprehensive quality assessment.

RELIABILITY

Reliability is the extent to which the results can be reproduced or are consistently reported. It is a measure of agreement between different abstractors (inter-rater reliability [IRR]) or multiple abstractions of the same record by the same abstractor (intra-rater reliability).

A common method of reporting results is % agreement. For IRR, many organizations consider 95% as the gold standard.[4,5]

ACCURACY

Accuracy refers to whether the value is the correct value or reflects the true state of the patient. It is measured by how closely the registry data match the source (EHR) data.

= number of correct fields / number of fields assessed

The North American Association of Central Cancer Registries (NAACCR) requires 100% accuracy for all the data variables used to create incidence statistics by cancer type, sex, race, age, and county for the gold-level certification and 97% accuracy for the silver-level certification.[6]

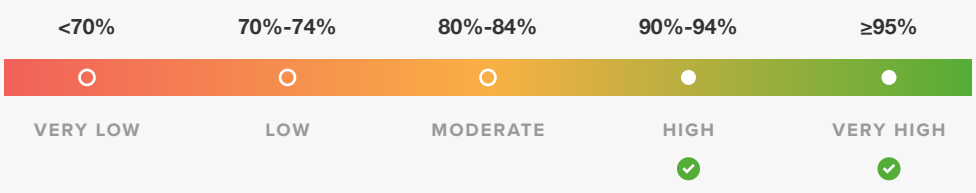
COMPLETENESS

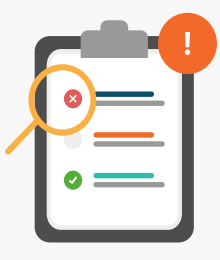
Completeness measures whether all necessary data have been entered. It is measured as:

= number of completed fields / number of fields assessed

For the gold-level NAACCR certification, at least 95% completeness is required, and at least 90% is required for the silver-level certification.[6]

Other registries have reported the following standards:[7]





VALIDITY

Validation is the process of determining whether the information falls within the acceptable range for the specific field.

This can be achieved by using automated range checks, for example to check that serum creatinine falls within the defined range of 15 mol/L to 600 mol/L.

Consistency checks can identify potential errors by checking if entered values fall outside the range in relation to other database variables. For example, the discharge date should be later than the admission date.

All measurements play an important role in assessing data quality.

At Carta Healthcare, we believe that all aspects of data quality are important for a quality assessment program. Each measurement has a role, and using only one or two measures does not adequately capture the inherent multidimensionality of the data and limits the utility of quality assessments. Further, these factors are related. For example, a training program that improves accuracy might decrease intra-rater reliability because after training, the abstractor will be more likely to choose the correct value that will not match the incorrect, pre-training value.

A robust quality assessment program therefore includes consistent use of these metrics: validation process ensures the plausibility of values, completeness checks identify blank or incomplete fields, IRR assessments disclose inconsistencies, and this feedback signals the need for additional review of the source record for accuracy.

Learn more about these measurements in our blog post.

WWW.CARTA.HEALTHCARE HELLO@CARTA.HEALTHCARE +1 (650) 338-4241

REFERENCES

1. Weiskopf NG, Weng C. Methods and dimensions of electronic health record data quality assessment: enabling reuse for clinical research. J Am Med Inform Assoc 2013;20:144-151.
2. Watzlaf VJM, Sheridan PT, Alzu'bi AA, et al. Clinical Data Abstraction: A Research Study. Perspectives in Health Information Management 2021.
3. Blumenthal S. The Use of Clinical Registries in the United States: A Landscape Survey. eGEMs (Generating Evidence & Methods to improve patient outcomes) 2017;5(1):1-13.
4. Liddy C, Wiens M, Hogg W. Methods to Achieve High Interrater Reliability in Data Collection From Primary Care Medical Records. Annals of Family Medicine 2011;9(1):57-62.
5. Cook EA, Schneider KM, Robinson J. Field methods in medical record abstraction: assessing the properties of comparative effectiveness estimates. BMC Health Services Research 2014;14:391-403.
6. North American Association of Central Cancer Registries. Certification Criteria. Available at: https://www.naaccr.org/certification-criteria/. Accessed on January 2, 2022.
7. Tan AC, Armstrong E, Close J, et al. Data quality audit of a clinical quality registry: a generic framework and case study of the Australian and New Zealand Hip Fracture Registry. BMJ Open Quality 2019;8:e000490. doi:10.1136/bmjoq-2018-000490