

# FAIR and Quality Assured Data -The Use Case of Trueness

Jürgen STAUSBERG<sup>a,1</sup>, Sonja HARKENER<sup>a</sup>, Ekkehart JENETZKY<sup>b,c</sup>,  
Patrick JERSCH<sup>d</sup>, David MARTIN<sup>b,e</sup>, Rüdiger RUPP<sup>d</sup> and Martin SCHÖNTHALER<sup>f</sup>  
<sup>a</sup>University Duisburg-Essen, Faculty of Medicine, Institute for Medical Informatics,  
Biometry and Epidemiology (IMIBE), Essen, Germany  
<sup>b</sup>Faculty of Health/School of Medicine, Witten/Herdecke University, Witten, Germany  
<sup>c</sup>Department of Child and Adolescent Psychiatry and Psychotherapy, University  
Medical Center of the Johannes-Gutenberg-University, Mainz, Germany  
<sup>d</sup>Heidelberg University Hospital - Spinal Cord Injury Center, Heidelberg, Germany  
<sup>e</sup>Department of Pediatrics, Eberhard-Karls University Tübingen, Tübingen, Germany  
<sup>f</sup>Department of Urology, Medical Centre - University of Freiburg, Faculty of Medicine,  
University of Freiburg, Freiburg, Germany

**Abstract.** The FAIR Guiding Principles do not address the quality of data and metadata. Therefore, data collections could be FAIR but useless. In a funding initiative of registries for health services research, trueness of data received special attention. Completeness in the definition of recall was selected to represent this dimension in a cross-registry benchmarking. The first analyses of completeness revealed a diversity of its implementation. No registry was able to present results exactly as requested in a guideline on data quality. Two registries switched to a source data verification as alternative, the three others downsized to the dimension integrity. The experiences underline that the achievement of appropriate data quality is a matter of costs and resources, whereas the current Guiding Principles quote for a transparent culture regarding data and metadata. We propose the extension to FAIR-Q, data collections should not only be findable, accessible, interoperable, and reusable, but also quality assured.

**Keywords.** Completeness, data quality, health services research, registries, validity.

## 1. Introduction

The FAIR Guiding Principles for scientific data management and stewardship [1] had been published in view of the scientific use of already recorded data. Data collections should be findable, accessible, interoperable, and reusable. Beside legal constraints, achieving the FAIR Guiding Principles is mainly a matter of provision and culture. The principles especially demand the willingness to share detailed information about a data collection with the public, a demand that might jeopardize economic and scientific gains of the primary data holder. However, even if a data collection fulfills the FAIR Guiding Principles, the data collection could be useless, because it does not offer a quality of data

---

<sup>1</sup> Corresponding Author, Institute for Medical Informatics, Biometry and Epidemiology, Faculty of Medicine, University Duisburg-Essen, Hufelandstrasse 55, 45122 Essen, Germany; E-mail: stausberg@ekmed.de.

needed to support the use case at hand. Furthermore, the re-use of data with inappropriate data quality can establish a hazard. In electronic health records, a wrong lab value such as an erroneously low serum potassium concentration - caused by a mix-up of patient identities - can lead to harm via inappropriate therapy. A lack in recording of adverse events in clinical trials could mask the risk of a drug therapy that is then misleadingly recommended in a clinical guideline. In the following, we will report experiences regarding the measurement of trueness from a funding initiative of registries for health services research.

## 2. Methods

### 2.1. *Quality Indicators Related to Trueness*

The guideline on data quality in medical research [2] organized indicators within three dimensions derived from Donabedian's approach: integrity, organization, and trueness. High quality in the first two dimensions is a necessary, but not a sufficient precondition to achieve a high quality in the dimension of trueness. The guideline defined six indicators for trueness: agreement with source data referring to data elements (TMF-1044), agreement with source data referring to observational units (TMF-1045), compliance with operating procedures (TMF-1047), representativeness (TMF-1048), accuracy (TMF-1043), and completeness (TMF-1046). Two indicators (1044, 1045) correspond with a source data verification well established in clinical trials. The assessment of the compliance with operating procedures requires the existence of such standards. Representativeness checks the agreement of frequencies and distributions from data with expectations, e.g. published in the literature. Accuracy and completeness capture precision and recall: did an event recorded in the data occur in reality (precision), was a real event recorded in the data (completeness)? Many proposals were made to operationalize those aspects also denoted as validity [3].

### 2.2. *Sample*

In a funding initiative, five indicators from the guideline [4] were selected for a benchmarking of registries. Trueness was represented by indicator completeness (TMF-1046). According to Arts et al. [5], completeness was defined as "the extent to which all necessary data that could have been registered have actually been registered". The calculation method of the indicator was adapted to each registry. This adaptation became part of the standard operating procedure (SOP) for the benchmarking of the registries. For example, the SOP listed the particular data elements considered for the measurement of completeness as well as the individual approach used to identify the clinical events that "could have been registered" in order to calculate the individual denominator. The quality target was adapted from the guideline as a completeness of 95% or higher.

The benchmarking of trueness for the year 2020 comprised five registries dealing with 1) fever in childhood, 2) lifelong monitoring of patients with spinal cord injury, 3) treatment exit options for uveitis, 4) hereditary breast and ovarian cancer, and 5) safety of living kidney donors. The indicator completeness was calculated by each registry itself accordingly to the SOP. An accompanying project was responsible for the receipt of the results and the preparation of the benchmarking report. The responsibility of the accompanying project did not include any monitoring or data management.

### 3. Results

The five registries delivered their results for the indicator completeness until February 2021. The determination of the rate's numerator and denominator was up to the registries. In comparison with the guideline's definition, the deliverables were quite diverse.

- Registry A reported a completeness of 112% for pre-existing-illnesses consulting medical reports.
- Registry B compared the relatively frequencies of three subtypes from the qualifying disease with expectations from two observational studies. With one exception, the 95% confidence limits included the literature results.
- Registry C counted adverse events that lead to a re-admission. The result was compared with secondary data. The calculated standardized ratio was 3.5.
- Related to the observational unit, but not controlled for the single event, registry D reported a completeness of 53% (contacts), 62% (drugs) and 87% (symptom).
- Registry E compared the recorded calendar date of an event with the calendar date of the event available in the patient chart. It found no discrepancies.

Based on these deliverables, the accompanying project did not prepare a benchmarking report for this indicator as specified in the SOP. Instead, the results were described and commented in the report without a comparison between the registries.

### 4. Discussion

Two out of the five registries achieved the intended data quality with a completeness of 95% or higher. However, none of the registries was able to calculate the indicator as specified in the guideline. The measurement of trueness seemed to be a big challenge in the monitoring of registries' data quality! Taking into account the whole set of indicators defined in the guideline [6], the registries made use of some other measures.

- Registry A performed a source data verification as specified in TMF-1044.
- Registry B assessed the evidence of known correlations (TMF-1027).
- Registry C also assessed the evidence of known correlations (TMF-1027).
- Registry D analyzed the concordance between two different data collections covering the same observational units (indicator TMF-1002).
- Registry E performed a source data verification as specified in TMF-1044.

Only two registries switched to another indicator from the dimension trueness, three fell back to the dimension integrity. Indicators from that dimension could be calculated without considering the origin and the context of data thus making data management substantially easier and cheaper. Considering external comparators is at least time- and labor-consuming. Furthermore, the establishment of a realistic gold standard could be expensive by creating organizational workflows in health care, solely related to this task.

Registry-based research is confronted with serious challenges. A main concern is related to a not-verified validity of important clinical events. The judgement of clinical outcomes is up to directly involved physicians [7], patients [8] or carers. In clinical trials, regulatory bodies recommend the establishment of a data monitoring committee [9], among others responsible for the verification of serious adverse events or outcomes. The recording of clinical events in daily practice is biased by the workload as well as the motivation of the responsible staff.

Precisely because the standards of clinical trials are out of the scope of observational research, data management of registries must have a clear understanding of alternative options to assure trueness. Medicine will not benefit from data collections fulfilling the FAIR Guiding Principles on the one hand but delivering an unsubstantiated data quality on the other hand. Therefore, we call for an extension of FAIR to FAIR-Q. In contrary to the already included requirements, to provide Quality assured data is a matter of resources, efforts and procedures. There must be resources for efforts implementing procedures of a quality oriented data management. The presented use case of trueness demonstrated that there is still a long way to achieve this goal. In analogy to the original proposal [1], “Q” can be characterized by requirements for data collections:

Q1. metadata are annotated regarding quality management procedures

Q2. indicators are available about the quality of (meta)data

Q2.1 a statement on the trueness of data is present

Q3. (meta)data are released with evidence about their potential impact

In our initiative, the registries could start with a weak-point analysis and subsequently with a redesign of their processes in case of insufficient results. The accompanying project will readjust the representation of the trueness to improve the cross-registry benchmarking. Two preferred options exist, to replace the indicator completeness with another indicator of that dimension or to rethink its specification for each registry. The calculation of the indicator by the registries themselves is a limitation of the benchmarking. So far, reliability and validity of the results remain open.

## Acknowledgements

The work was funded by the German Federal Ministry of Education and Research within the "Model Registers Funding Scheme" (FeverApp 01GY1905, ParaReg 01GY1904, RECUR 01GY1902, REGISVF-APB 01GY1917B).

## References

- [1] Wilkinson MD, Dumontier M, Aalbersberg IJ, et al. The FAIR Guiding Principles for scientific data management and stewardship. *Sci Data*. 2016;3:160018.
- [2] Nonnemacher M, Weiland D, Neuhäuser M, Stausberg J. Adaptive management of data quality in cohort studies and registers: proposal for a guideline. *Acta Informatica Medica*. 2007;15:225-30.
- [3] Stausberg J, Kuklik N, Jöckel KH. Opportunities and Pitfalls in the Definition of Data Validity. *Stud Health Technol Inform*. 2018;247:566-70.
- [4] Stausberg J, Bauer U, Nasseh D, et al. Indicators of data quality: review and requirements from the perspective of networked medical research. *GMS Med Inform Biom Epidemiol*. 2019;15:Doc05.
- [5] Arts DG, De Keizer NF, Scheffer GJ. Defining and improving data quality in medical registries: a literature review, case study, and generic framework. *J Am Med Inform Assoc*. 2002;9:600-11.
- [6] Harkener S, Stausberg J, Hagel C, Siddiqui R. Towards a core set of indicators for data quality of registries. *Stud Health Technol Inform*. 2019;267:39-45.
- [7] Roovers J-PWR. Registries: what level of evidence do they provide? *Int Urogynecol J*. 2007;18:1119–20.
- [8] Martin D, Wachtmeister J, Ludwigs K, Jenetzky E. The FeverApp registry - ecological momentary assessment (EMA) of fever management in families regarding conformity to up-to-date recommendations. *BMC Med Inform Decis Mak*. 2020;20:249.
- [9] U.S. Department of Health and Human Services, Food and Drug Administration. Guidance for Clinical Trial Sponsors. Establishment and Operation of Clinical Trial Data Monitoring Committees. March 2006. <https://www.fda.gov/media/75398/download> (access 2021-06-21).