Assessing the Quality of Real-World Data and Real-World Evidence in Oncology Research: A Cohesive Framework for Researchers

ontada

The US Oncology Network

Zhaohui Su, PhD, Joseph T. Dye, PhD, RPh, Thomas Wilson, PhD, DrPH, E. Susan Amirian, PhD , Amy K. O'Sullivan, PhD

Objectives

- Summarize the quality domains in existing frameworks
- Propose a new tool for assessing the quality of real-world data (RWD) and real-world evidence (RWE) for oncology studies

Methods

- A literature review was conducted to identify methods for determining the quality of RWD and RWE.
- Challenges in oncology research were discussed among scientific experts.
- We developed a new method that assesses the quality of RWD and RWE from a broader perspective, with the goal to help understand whether and what oncology research questions can be answered with the RWD available.
- The proposed method was based on 9 quality domains identified in the literature review.

Results

- We reviewed 10 frameworks and selected a total of 9 domains. Each framework includes between 3 and 6 data quality domains. Our proposed tool includes all 9 domains.
- The proposed method assesses quality of RWD and RWE from a broader perspective and takes into consideration the unique challenges in oncology research. For example, scalability is important for oncology research because many data are only available in patient's medical charts and are not available in structured fields.
- Each assessment of the 9 quality domains returns a binary YES/NO result. These assessments are advised to be assessed under relevant published guidance with well-documented operational definitions and business rules.
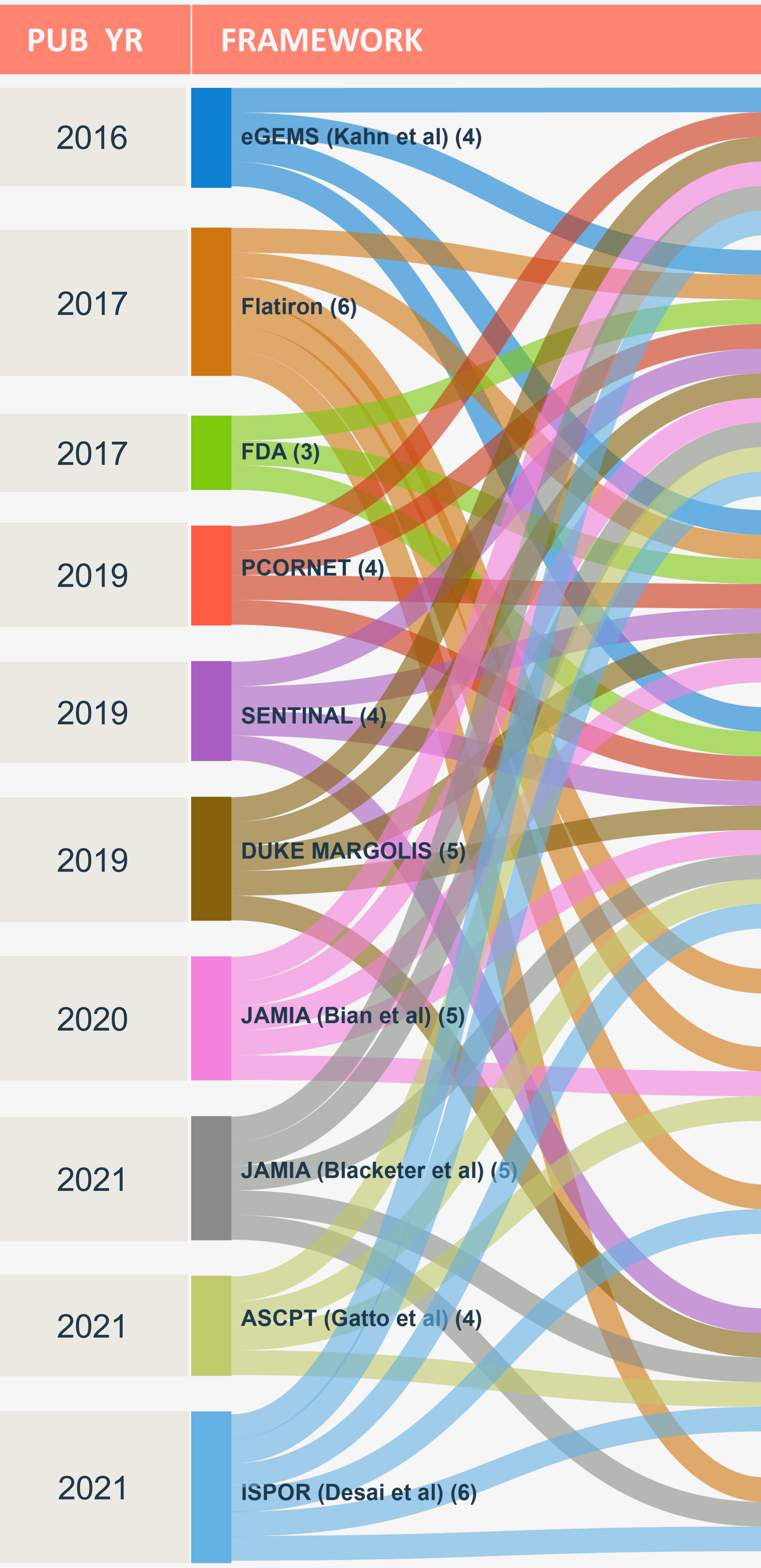
Figure 1. Published frameworks and quality domains

Figure 1. Published frameworks and quality domains

PUB YR	FRAMEWORK
2016	eGEMS (Kahn et al) (4)
2017	Flatiron (6)
2017	FDA (3)
2019	PCORNET (4)
2019	SENTINAL (4)
2019	DUKE MARGOLIS (5)
2020	JAMIA (Bian et al) (5)
2021	JAMIA (Blacketer et al) (5)
2021	ASCPT (Gatto et al) (4)
2021	ISPOR (Desai et al) (6)

Table 1. Assessment questions by quality domain

DOMAIN	QUESTION (YES/NO)	RESEARCH EXAMPLE
Conformance (6)	Are data compliant to dictionary or specifications?	Race categories are not collected per data dictionary.
Completeness (10)	Are data values present?	Gender should not be null.
Consistency (7)	Are data being collected the same way?	Data are collected in different ways across practices.
Accuracy (9)	Are data values believable?	Death dates can't be before birth dates.
Scalability (1)	Can large and complicated data be processed efficiently?	Data may be stored as structured data, or in notes, or as pictures.
Timeliness (3)	Are data recent and timely?	Data are refreshed daily or weekly.
Generalizability (2)	Is RWE based on a broad range of representative patients (i.e., external validity)?	Death dates can't be before birth dates.
Validity (5)	Is the relationship being tested trustworthy (i.e., internal validity)?	Potential biases, confounding and missing data have been handled properly.
Transparency (3)	Are study design and planned analyses described in detail?	Ad hoc analyses need to be carefully documented.

Table 2. Proposed framework

Proposed RWD quality rating	High quality if having 6 "YES"s to first 6 domain questions, moderate quality if having at least 4 "YES"s to the first 6 domain questions, and low quality otherwise.
Proposed RWE quality rating	High quality if having 3 "YES"s to all last 3 domain questions, moderate quality if having at least 2 "YES"s to the last 3 domain questions, and low quality otherwise.
Proposed RWD/RWE (overall) quality rating	High quality if having 9 "YES"s to the 9 domain questions, moderate quality if having at least 7 "YES"s to the 9 domain questions, and low quality otherwise.

Conclusion

- We developed a new tool to provide a more comprehensive approach for assessing RWD and RWE quality. It was built off 9 domains from existing frameworks.
- The new tool has 3 ratings: RWD quality, RWE quality, and an overall quality rating.
- The new tool will aid in assessing fit-for-purpose, and will broadly help researchers understand the RWD and RWE quality more broadly and for specific domains.
- This has important implications from a data ownership and data license perspective, and for generating RWE to understand the natural history of disease and the effectiveness of medical products.
- More validation work is ongoing.
- This tool can be expanded and used for other disease conditions.

Scan here to learn more