Original Paper

Implementing Findable, Accessible, Interoperable, Reusable (FAIR) Principles in Child and Adolescent Mental Health Research: Mixed Methods Approach

Rowdy de Groot^{1,2}, MSc; Frank van der Graaff³; Daniël van der Doelen⁴, MSc; Michiel Luijten^{5,6,7}, PhD; Ronald De Meyer³, PhD; Hekmat Alrouh⁸, MSc; Hedy van Oers^{5,7}, PhD; Jacintha Tieskens⁹, PhD; Josjan Zijlmans^{7,10}, PhD; Meike Bartels^{7,8}, PhD; Arne Popma^{7,10,11}, PhD; Nicolette de Keizer^{1,2}, PhD; Ronald Cornet^{1,2*}, PhD; Tinca J C Polderman^{4,7,9,10,11,12*}, PhD

³Praktikon, Nijmegen, Netherlands

⁶Amsterdam University Medical Center, Vrije Universiteit Amsterdam, Epidemiology and Data Science, Amsterdam, Netherlands

⁷Amsterdam Public Health, Mental Health, Amsterdam University Medical Center, Amsterdam, Netherlands

⁸Department of Biological Psychology, Vrije Universiteit Amsterdam, Amsterdam, Netherlands

⁹Leiden University Medical Center Curium–Child and Adolescent Psychiatry, Leiden University Medical Center, Leiden, Netherlands

¹⁰Department of Child and Adolescent Psychiatry & Psychosocial Care, Amsterdam University Medical Center, Vrije Universiteit Amsterdam,

Amsterdam, Netherlands

¹¹Levvel, Academic Center for Child and Adolescent Psychiatry, Amsterdam, Netherlands

Corresponding Author:

Rowdy de Groot, MSc Department of Medical Informatics Amsterdam University Medical Center, University of Amsterdam Meibergdreef 9 Amsterdam, 1105 AZ Netherlands Phone: 31 648499049 Email: rowdy.degroot@amsterdamumc.nl

Abstract

Background: The FAIR (Findable, Accessible, Interoperable, Reusable) data principles are a guideline to improve the reusability of data. However, properly implementing these principles is challenging due to a wide range of barriers.

Objectives: To further the field of FAIR data, this study aimed to systematically identify barriers regarding implementing the FAIR principles in the area of child and adolescent mental health research, define the most challenging barriers, and provide recommendations for these barriers.

Methods: Three sources were used as input to identify barriers: (1) evaluation of the implementation process of the Observational Medical Outcomes Partnership Common Data Model by 3 data managers; (2) interviews with experts on mental health research, reusable health data, and data quality; and (3) a rapid literature review. All barriers were categorized according to type as described previously, the affected FAIR principle, a category to add detail about the origin of the barrier, and whether a barrier was mental health specific. The barriers were assessed and ranked on impact with the data managers using the Delphi method.

Results: Thirteen barriers were identified by the data managers, 7 were identified by the experts, and 30 barriers were extracted from the literature. This resulted in 45 unique barriers. The characteristics that were most assigned to the barriers were, respectively, external type (n=32/45; eg, organizational policy preventing the use of required software), tooling category

¹Department of Medical Informatics, Amsterdam University Medical Center, University of Amsterdam, Amsterdam, Netherlands

²Amsterdam Public Health, Digital Health, Amsterdam University Medical Center, Amsterdam, Netherlands

⁴Karakter Child and Adolescent Psychiatry University Centre, Nijmegen, Netherlands

⁵Emma Children's Hospital, Child and Adolescent Psychiatry & Psychosocial Care, Amsterdam University Medical Center, University of Amsterdam, Amsterdam, Netherlands

¹²Department of Child and Adolescent Psychiatry, University Medical Center Groningen, University of Groningen, Groningen, Netherlands *these authors contributed equally

(n=19/45; ie, software and databases), all FAIR principles (n=15/45), and not mental health specific (n=43/45). Consensus on ranking the scores of the barriers was reached after 2 rounds of the Delphi method. The most important recommendations to overcome the barriers are adding a FAIR data steward to the research team, accessible step-by-step guides, and ensuring sustainable funding for the implementation and long-term use of FAIR data.

Conclusions: By systematically listing these barriers and providing recommendations, we intend to enhance the awareness of researchers and grant providers that making data FAIR demands specific expertise, available tooling, and proper investments.

JMIR Ment Health 2024;11:e59113; doi: 10.2196/59113

Keywords: FAIR data; research data management; data interoperability; data standardization; OMOP CDM; implementation; health data; data quality; FAIR principles

Introduction

The FAIR (Findable, Accessible, Interoperable, Reusable) data principles [1], published in 2016, are a guideline to improve the reusability of health data. These principles imply, respectively, the following: (1) Findable—metadata (ie, data about the data) and data should be easy to find for both humans and computers; (2) Accessible—the user needs to know how data and metadata can be accessed, possibly including authentication and authorization; (3) Interoperable—the data usually need to be integrated with other data; the data need to interoperate with applications or workflows for analysis, storage, and processing; and (4) Reusable—metadata and data should be well described so that they can be replicated and combined in different settings [2].

The implementation of the FAIR principles and thereby the increase of availability and usability of data advances scientific knowledge in several ways. First, it benefits researchers who have no sources for data collection themselves, for example, from low- and middle-income countries. Second, it enhances data sharing across cohorts, registers, or study samples worldwide, thus enlarging the scale, statistical power, and impact of empirical studies. Finally, it invites researchers from different disciplines to provide their unique perspectives and expertise on the exchanged data. In short, FAIR data maximize the usability of carefully collected, valuable data. In addition, FAIR data can contribute to a clear overview of which data are available for specific research questions (and as such avoid spending resources on collecting similar data) and facilitate the potential replication of published research, all adding to the reliability, validity, and value of scientific knowledge in general. Researchers increasingly aim to implement the FAIR principles [3]. In a study by Kersloot et al [4], after reading a short description of the FAIR principles, scientists were asked how much effort they spent on making their data more FAIR. About 80% indicated that they have spent effort to make their data more FAIR, addressing 1 or more of the 4 aspects [4].

However, despite a positive attitude of researchers regarding the FAIR principles, it appears challenging to properly implement them [5,6]. The FAIR principles provide guidance on *what* to do but not on *how* to do it. Implementing the FAIR principles requires expert knowledge of data, metadata, identifiers, ontologies, and terminologies, but there is a lack of guidelines and tooling, leaving it to researchers to assess which technology or resources should be used to achieve FAIRness.

For health data, it might be even more difficult to implement the FAIR principles, since one of the considerable limitations is privacy regulation, such as the European Union's General Data Protection Regulation (GDPR). This complicates the "A"—accessibility aspect of FAIR, as more complex access procedures need to be described due to required security measurements. Of note, FAIR is often confused with open data; however, this is not the same. FAIR aims to facilitate federation, which deals (at least in part) with privacy matters and does not aim to make data freely available to everyone.

This study focuses on the FAIRification of mental health data. More than 13% of people worldwide suffer from mental illness [7], which has a major impact on global health. Thus, optimizing knowledge development in this health domain is crucial [8,9]. However, Sadeh et al [9] conclude that the mental health domain has not yet developed a culture around FAIR data or required tools and resources to easily FAIRify mental health data. We aim to identify the concrete barriers to FAIRification in the domain of mental health data and provide recommendations for these barriers to move the field of FAIR data in general and specifically for mental health further.

Methods

Overview

We used three sources as input to identify barriers: (1) evaluation of the implementation process of the Observational Medical Outcomes Partnership (OMOP) Common Data Model (CDM) by 3 data managers at 3 research sites; (2) interviews with experts in mental health research, reusable health data, and data quality; and (3) a rapid literature review. The research was directed by a qualified data steward (RdG), who categorized the identified barriers. Using the Delphi method [10], the 3 data managers scored the impact of the barriers on the FAIRification process. Subsequently, the top 10 most impactful barriers were determined. Recommendations for these barriers were derived from the authors and relevant scientific literature based on the rapid review.

OMOP CDM

OMOP CDM is an open community data standard maintained by the OHDSI (Observational Health Data Sciences and Informatics) community that is used to represent observational data in a standardized manner. This standard ensures both semantic interoperability, where data are understood across systems, and syntactic interoperability, where the format of data is usable across systems. This allows researchers to write programs or analyses that work on multiple OMOP CDM instances. It also allows researchers to combine data, although OMOP CDM is more focused on federated analyses (ie, researcher A writes a script that researcher B can run on their OMOP CDM instance and researcher B sends back the results, thus without sharing data). OMOP CDM makes it also possible to use the OHDSI research tools and to perform collaborative research or large-scale analytics [11].

A data standard to increase the interoperability of the collected data was required. Therefore, together with the data managers we chose to implement OMOP CDM since it is oriented toward health-related observational data and makes the data more interoperable (the I of FAIR).

Research Sites

The 3 research sites that participated in implementing OMOP CDM were KLIK [12], DREAMS (Dutch REsearch in child and Adolescent Mental health), and the Learning Database Youth (in Dutch: LDJ) [13,14]. These 3 research sites, among other research sites, were involved in a consortium to assess the mental health problems in Dutch children and adolescents

before and during the COVID-19 pandemic [15,16]. These sites used a variety of questionnaires to collect data on mental health and demographic variables, such as sex, age, and socioeconomic status, from children, adolescents, and their parents with and with no preexisting mental health problems. These 3 research sites were part of a national consortium that collaborated on research on mental health of children and adolescents during the COVID-19 pandemic and concerned all sites that intended to harmonize data in the participating research sites.

de Groot et al

KLIK is established in an academic hospital and collects data through a dedicated research website for research and to monitor children. DREAMS is a collaboration between 4 academic child and adolescent psychiatry centers that collects data for research and aims to improve mental health care for children. LDJ is a cooperation between more than 30 youth care institutions in the Netherlands that collects data on the mental health of children and adolescents receiving youth care to improve the quality of care. An overview of the research sites and their data represented in OMOP is shown in Table 1. Each of the 3 research sites had a data manager that was tasked to implement OMOP CDM (DvdD, FvdG, and ML). The 3 data managers were supported by a data steward experienced in implementing OMOP CDM (RdG). None of the research sites or data managers had previous experience with FAIR data or the implementation of OMOP CDM. Two of the 3 data managers had no experience with Structured Query Language, a standardized programming language that is used to manage relational databases and is necessary for the implementation of OMOP CDM.

Table 1. Overview of research sites and datasets (the Cantrill Ladder score is used to measure well-being).

	KLIK	DREAMS ^a	LDJ ^b
Description	Developed in an academic hospital and collects data through a research website developed for collecting data	A collaboration between 4 academic child and adolescent psychiatry centers, collects data for research, and aims to improve mental health care for children	A cooperation between more than 30 youth care institutions in the Nether- lands that collects data on the mental health of children and adolescents receiving youth care to improve the quality of care
Data subset implemented in OMOP ^c	Person ID, age, sex, Cantrill Ladder score, PSQ ^d , and PROMIS ^e anxiety	Person ID, sex, Cantrill Ladder score, PSQ, PROMIS anxiety, and date of questionnaire	Person ID, date of birth, sex, PSQ, and date of questionnaire
aDDEAMS, Dutch DEcourch in a	hild and Adalasaant Mantal haalth		

^aDREAMS: Dutch REsearch in child and Adolescent Mental health.

^bLDJ: Learning Database Youth.

^cOMOP CDM: Observational Medical Outcomes Partnership Common Data Model.

^dPSQ: Parenting Stress Questionnaire.

^ePROMIS: patient-reported outcomes measurement information system.

Barriers

In this paper, we define a barrier as an impediment to the implementation of the FAIR data principle. Usually, a barrier requires more resources, people, funding, or time to solve. For example, the GDPR does not prevent the implementation of the FAIR principles but requires more knowledge on what is and is not allowed. This could lead to the need for technical measures, for example, to prevent unrestricted access to the data.

First, barriers were identified by 3 Dutch data managers (FvdG, DvdD, and ML) during the implementation of OMOP

CDM [11] for data from 3 mental health research sites, as described in the "Research sites" section. Second, we also collected barriers mentioned during interviews with a team of experts on reusable health data (RC), child and adolescent behavioral data (TJCP), and data quality (NdK). Third, we searched the scientific literature for relevant studies on implementing the FAIR principles and potential barriers. The process of collecting barriers and determining the results is shown in Figure 1.

Figure 1. Overview of the process of the study. The entire process was coordinated by the data steward (RdG). Tooling: ATLAS is a tool to create cohorts in OMOP CDM; Cantrill Ladder score is used to measure well-being; Data Quality Dashboard is used to assess the quality of the data in OMOP CDM; Rabbit-in-a-Hat is a tool to design the conversion to OMOP CDM; Usagi is a tool to annotate local terms to standard concepts in OMOP CDM; White Rabbit is a tool to scan data. DREAMS: Dutch REsearch in child and Adolescent Mental health; ETL: Extract Transform Load; LDJ: Learning Database Youth; OMOP CDM: Observational Medical Outcomes Partnership Common Data Model; PROMIS: patient-reported outcomes measurement information system; PSQ: Parenting Stress Questionnaire.



OMOP CDM Implementation

Since none of the data managers had previous experience with OMOP CDM, and the research centers collected partly different data elements, we started by modeling in OMOP CDM a small subset of the research data of each site. The subset consisted of a person ID, age or date of birth, sex, a Cantrill Ladder assessment of well-being [17], a Parenting Stress Questionnaire (PSQ) [18], a PROMIS (patient-reported outcomes measurement information system) questionnaire assessing anxiety [19], and the dates of filling in the questionnaires. For the DREAMS and KLIK research sites all data items were available while the LDJ research site did not collect the Cantrill Ladder assessment and the PROMIS anxiety questionnaire. Representative synthetic data were used in the OMOP CDM implementations.

We followed standard OHDSI steps for the OMOP CDM implementation process [20]. The OMOP process started by analyzing the data with White Rabbit. White Rabbit is a tool provided by OHDSI to scan datasets to provide information on tables, fields, and values. White Rabbit generates a scan report that is necessary to use Rabbit-in-a-Hat. OHDSI provides Rabbit-in-a-Hat to help users design their Extract, Transform, Load (ETL) process to represent their data in OMOP CDM. Usagi was used by the DREAMS and KLIK research sites to annotate local data element names to standardized concepts from the OMOP vocabulary for the Cantrill Ladder assessment and the PROMIS pediatric anxiety. The next step was to implement the ETL in a database and check the quality of the ETL with the data quality dashboard [21]. The last step was to use Atlas, a tool to conduct (research) analyses on OMOP CDM data [22], to construct a cohort and show that analyses can be performed on the converted OMOP CDM data. During the implementation process, 15 designated meetings were planned in which the data steward discussed the progress of the OMOP process and the encountered barriers with the data managers.

Expert Interviews

The data steward (RdG) had expert interviews with 3 experts involved in the research from 3 different perspectives.

- 1. RC is a professor of medical informatics, focusing on semantic interoperability, both from a technical perspective and from a users' point of view, as a key component to establishing FAIR data for health care and research.
- 2. TJCP is an associate professor of psychology and a research expert on child and adolescent behavioral data.
- 3. NdK is a professor of medical informatics, specializing in evaluating health care and health care information systems.

The interviews were held during 2 designated meetings in which the barriers identified by the data managers were reviewed by the experts. The experts together with the data steward (RdG) reflected on the barriers and complemented these with specific barriers from the experts. One additional meeting was held with a group of experts in child and

https://mental.jmir.org/2024/1/e59113

adolescent mental health care during which the barriers from the data managers and experts were discussed.

Literature

A rapid literature review was performed to identify barriers that hamper the implementation of the FAIR principles to both confirm and complement those encountered during the OMOP process or discussed during the expert interviews. The search term was "FAIR data implementation" in Google Scholar. No explicit inclusion or exclusion criteria were defined. If the title and abstract mentioned FAIR implementation barriers, we extracted challenges and barriers from the full text. We focused on FAIR data implementations in general to include as many papers as possible to make sure that we included a wide range of barriers. We searched only in Google Scholar since Google Scholar also encompasses results from other search engines such as PubMed.

Categorizing the Barriers

All barriers to implementing the FAIR principles were categorized by the data steward. The first categorization was the type of barrier as described by Cabana et al [23]. The types of barriers were "Lack of Awareness," "Lack of Familiarity," "Lack of Agreement," "Lack of Self-efficacy," "Lack of Outcome Expectancy," "Inertia of Previous Practice," and "External Barriers." Each barrier could be of 1 or more types of barriers. The second categorization regarded the FAIR principles or aspects of FAIR that were hampered due to the barrier. To further detail the origin of the barrier, a third categorization (the category) was added in which we specified whether people (skills or knowledge), tools, material, process, or costs were involved. Finally, we specified whether or not barriers were specific to the mental health care domain.

Modified Delphi Method

We conducted a modified Delphi method as defined by Boulkedid et al [10] to establish consensus on the most impeding barriers to implementing the FAIR principles. First, a questionnaire including all identified barriers was sent to the 3 data managers to ask them to independently rate the barriers on a 1-to-10 scale on how much a barrier would hinder the implementation of the FAIR principles if it was encountered. The scores were totaled to determine the final score. The second round was a web-based meeting where the results were discussed with the supervision of the data steward. Barriers that had a scale score difference of 4 or more between any of the data managers were discussed, and clarified by the data steward, and scores were adjusted with justification if necessary. We considered a difference of 3 or fewer to reflect that each barrier could have a different experienced impact but still close enough for consensus. Justification of using the Delphi method is shown in Multimedia Appendix 1 according to the Guidance on Conducting and REporting DElphi Studies (CREDES) by Jünger et al [24]. We calculated percent agreement as a simple index for agreement between the 3 data managers and we calculated Krippendorff α as interrater reliability index [25]. Krippendorff α values range from 0 to 1, where 0 is

perfect disagreement and 1 is perfect agreement. An α value higher than .8 is recommended whereas .667 is the lowest conceivable limit [26].

Recommendations to Overcome the Barriers

Based on the results of the modified Delphi, we determined the top 10 most problematic barriers to focus on the most impactful barriers. We chose a top 10 list to provide a compact set of detailed recommendations, rather than a long list that would make it difficult to prioritize. In addition, some barriers, such as standardizing answers in open questions, cannot be solved. We believe that addressing the top 10 most problematic barriers will also help resolve many other issues, thus avoiding repetition. The barriers that scored highest were further examined to assess whether recommendations for overcoming these barriers were reported in the literature or could be determined or devised by the authors (RC and RdG, both experienced with FAIR data). We decided not to rank the recommendations, as they were virtually matched to all identified barriers. By ranking the barriers, we indirectly ranked the recommendations as well.

Ethical Considerations

This study did not involve medical research and did not regard any interventions; no ethics approval was applied for in this study in line with the ethics approval procedures described on the Amsterdam Universitair Medische Centra website [27].

Results

Barriers

In total, 45 unique barriers to implementing the FAIR principles were identified. Of these, 13 barriers were identified during the OMOP CDM implementation process, 7 barriers were identified by experts, and 30 barriers were found in literature [6,28-36]. Five barriers were identified by all 3 sources. Multimedia Appendix 2 shows all identified barriers and their sources.

Categorizing the Barriers

The "External barrier" type was assigned to most barriers: 71% (32/45). External barriers are barriers that hamper a researcher from implementing the FAIR principles, for example, organizational policy preventing the use of required software. Many barriers affected all of the FAIR principles (15/45, 33%). Interoperability and reusability principles were the most affected individual FAIR data aspects by 18% (8/45) and 20% (9/45) barriers, respectively. Principle I2: "(Meta)data use vocabularies that follow the FAIR principle" was the most affected FAIR principle with 9% (4/45) barriers, for example, answers to open questions in questionnaires cannot be standardized due to all answer possibilities. Approximately 42% (19/45) barriers were categorized in the

"tools" category which was the most prevalent category, for example, updates can break software dependencies. Only 4% (2/45) of barriers were specific to mental health, namely (1) mental health data are collected in care modules (subcare pathways) that contain only questionnaires, and (2) the *DSM* (*Diagnostic and Statistical Manual of Mental Disorders*, ie, one of the core diagnostic instruments in mental health) does not make use of identifiers. Identifiers are necessary to refer to relevant metadata of the diagnosis. Approximately 11% (5/45) of barriers are more likely to occur in the mental health domain, mostly due to the extensive use of questionnaires, yet not exclusively applicable to mental health. All categorizations and scores of the barriers are shown in Multimedia Appendix 2.

Modified Delphi Method

One data manager required more information from the data steward regarding certain barriers to score these. For 20 out of 45 barriers, consensus was reached by the data managers after filling in the questionnaire. The data managers reviewed the barriers for which there was no initial consensus, that is, a difference of 4 points or more between ratings. During the subsequent first live meeting, the 3 data managers discussed the 25 barriers without consensus after the questionnaire. During this meeting, for 21 out of 25 barriers the score was changed, resulting in a consensus for 19 barriers, while 6 barriers remained without reaching a consensus. The 6 barriers without consensus were again discussed during the second live consensus meeting in which the data steward participated in the discussion and consensus was reached for all barriers. The percentage of agreement was 55.6% before the first live consensus meeting, 86.7% after the first live consensus meeting, and 100% after the second live consensus meeting. Krippendorff α values were 0.34 (95%) CI 0.15-0.52), 0.67 (95% CI 0.47-0.81), and 0.79 (95% CI 0.70-0.85), respectively, of which only the last value was close to the recommended value of 0.8. The lowest scoring barrier (answers to open questions in questionnaires cannot be standardized) had a score of 4 and the highest scoring barrier (Lack of technical knowledge or access to people with the necessary knowledge) had a maximum score of 30. All notes, scores, and comments of each round are shown in Multimedia Appendix 3.

Recommendations to Overcome the Barriers

The recommendations for the top 10 barriers were derived from authors RC and RdG and the literature [37-42]. Table 2 shows the top 10 barriers and the recommendations to overcome these barriers. The main recommendations are to involve FAIR data stewards, create step-by-step guides for researchers to make their data FAIR, and implement sustainable funding that can maintain and support the FAIR data infrastructure, once it has been set up for a certain project.

de Groot et al

Table 2. The top 10 impeding barriers of implementing the FAIR^a principles and recommendations to solve these problems.

	impeding barrens of implementing the FAIR principles and	recommendations to solve these problems.
Number and score	Top 10 barriers	Recommendation
#1 (30 points)	Lack of technical knowledge (of programs, languages, and error messages) or access to people with the necessary knowledge.	(1a) Researchers need a step-by-step guide to FAIRify their data that is also easy to understand with little technical knowledge. This guide should also include a section to make resources reusable by computers. This guide does not need to be useable by everyone but could apply to a specific community or a single organization. (1b) A checklist of required tools and software might provide useful as well and could be combined with teaching handbooks such as the handbook written by Engelhardt et al [39]). Many barriers we identified are related to tooling. Tooling and proper documentation could potentially help out in overcoming a lack of technical knowledge. (1c) Another approach is involvement of a (FAIR) data steward who will support the researchers in correctly implementing the FAIR principles (Scholtens et al [37]; Wendelborn et al 2023 [38]).
#2 (28 points)	The infrastructure to host FAIR data must remain active/ online after the project or research is stopped.	Research funders need to provide a funding structure by which costs of making data FAIR and maintaining FAIR data infrastructure are covered, including the period after the project has ended or alternatives such as public or institutional repositories need to be considered to cover these costs.
#3 (28 points)	Organizational policies prevent researchers from implementing FAIR principles. Organizational support is required for setting up the infrastructure, including installing required software and databases.	Organizations need to have a streamlined process for their researchers to make their data FAIR and make sure that necessary tools and software are available for researchers (Jacobsen et al [40], FAIRification workflow, FAIR cookbook). See #1a. and #1b. for recommendations.
#4 (27 points)	Updates can break software dependencies.	Researchers can consider software that enables use of containerized applications such as Docker, Podman, or Buildah. However, the ability to implement software like this might depend on other barriers, such as technical skills or organizational policies that may prevent the implementation of certain software.
#5 (26 points)	Communication between different stakeholders that have different backgrounds is a challenge in a FAIRification process. A FAIR data steward seems essential to manage the necessary communication between disciplines (Queralt et al [6]).	Scholtens et al [37] argue that a (FAIR) data steward is vital for the communication between different disciplines and professionals (Scholtens et al [37]). We recommend that funding bodies or research-performing organizations have (FAIR) data stewards that can join projects and manage the research data management aspects of a project and can communicate this to project members.
#6 (26 points)	Stakeholders are not familiar with the steps that are needed to make a resource reusable by computers across multiple locations (Queralt et al [6]). Making data FAIR in a community without guides to make data FAIR is challenging.	See #1a for recommendations for step-by-step guides.
#7 (25 points)	Unclear who can or wants to handle data access requests especially after the project is stopped since funding is required to cover costs after the project is finished.	(Mental) health data are sensitive and will probably be available only upon a reasonable request. Therefore, someone needs to handle data access requests. The main problem of this barrier is the cost of someone that can handle the data access request after the project has ended. For funding recommendations, see #2.
#8 (25 points)	General Data Protection Regulation (GDPR) compliance needs to be embedded and the GDPR "right to be forgotten" should not be forgotten in the FAIR implementation plan (Overmars et al [29]).	Researchers need a clear guide on what the GDPR precisely entails. What is considered to be "personal data" can slightly differ between countries, indicating that there are still problems interpreting the GDPR. A (FAIR) data steward should be familiar with the GDPR and could provide help. See #1c. for FAIR data stewardship recommendations.
#9 (25 points)	Implementing the FAIR principles is a resource-intensive task, especially when carried out retrospectively. Resource costs associated with this task include the time, effort, and potentially standing up expenses (Alharbi et al [31]). FAIRifying data retrospectively remain challenging (Rocca-Serra and Sansone [43]).	For step-by-step guide and tooling recommendations see #1a. For funding recommendations, see #2.
#10 (25 points)	Legal challenges correspond to requirements that might pertain to the processing and sharing of the data (eg, accessibility rights and compliance with data protection regulations), both for meeting the "accessibility" and "reusability" criteria as well as for performing the	For legal challenges recommendations (eg GDPR), see #8.

Number and score	Top 10 barriers	Recommendation		
	FAIRification process itself (Wise et al 2022 [30]). The legal aspect of access rights is a significant issue due to its complexity and the lack of a clear process for accessing			
	previously generated datasets (Alharbi et al 2021 [31]). Protecting patient data and privacy, which could require restrictions (also mentioned by Queralt-Rosinach et al [6]).			
FAIR: Findable, Accessible, Interoperable, Reusable.				

Discussion

In this study, we identified 45 barriers that hamper the implementation of the FAIR principles in child and adolescent mental health research. Our study showed that most barriers are external barriers that have to do with tooling, and these mostly hamper all aspects of FAIR.

Although this research was conducted in the context of child and adolescent mental health, only 2 barriers were specific for mental health. One barrier is that mental-health data are collected in care modules that contain only questionnaires that made it difficult to represent in OMOP CDM and the other is that the DSM does not make use of identifiers. All other barriers were considered general barriers for making health data FAIR. The data in the child and adolescent mental health domain are therefore not more difficult to make FAIR than those of other health care domains. Moreover, the 2 mental health-specific barriers do not impact key components of the FAIR principles and are not in the top 10 barriers. One barrier would be solved by adding identifiers to the DSM-5 (Diagnostic and Statistical Manual of Mental Disorders [Fifth Edition]) diagnoses. The other barrier is more complex as it requires community-wide guidelines on how to represent care modules (subcare pathways) in data standards.

This research was conducted in the Netherlands. However, the results show no indication of FAIR barriers that are country-specific. Meaning, that initiatives such as the United States Core Data for Interoperability Plus (USCDI+) [44] could learn from other initiatives across countries without taking country-specific barriers into account.

The top 10 barriers that we identified show that researchers struggle to put the FAIR principles into practice. Based on our study, we have three recommendations to enhance FAIR data use: (1) the structural use of FAIR data stewards [45], (2) creating clear step-by-step guides to make data FAIR, and (3) implementing sustainable funding systems to implement and maintain the FAIR data infrastructure.

Research from Garcia et al [5], Queralt-Rosinach et al [6], van Vlijmen et al [33], Jacobsen et al [35], and Mayer et al [46] already mentions challenges of implementing the FAIR principles. By combining all barriers, we intend to provide a valuable overview of barriers for researchers and funders and help them to be proactive in implementing structural solutions that can help prevent these barriers from occurring.

The results of this study show that many barriers exist in the data FAIRification process. As such, they provide guidance for researchers to prepare themselves before starting the FAIRification process by assessing our overview of barriers and implementing the recommendations. We decided to use the term "recommendations" instead of "strategies" because our recommendations can also apply to factors not directly influenced by researchers, such as organizational policies. We wanted to clearly convey this distinction. Furthermore, these results are intended to support researchers to better communicate their needs to organizations or grant providers for making their data FAIR. A caveat is that the recommendations could introduce new barriers. For instance, in our study, a FAIR data steward was closely connected to the consortium. He attended research meetings, was involved in solving data problems, and helped implementing the FAIR principles. However, the introduction of a FAIR data steward brings costs, which could be a financial barrier. Furthermore, there is a need for further capacity building of FAIR data stewards.

A strength of this study is that we used 3 distinct sources to identify FAIR data implementation barriers: an extensive list of barriers was created by combining the experience of 3 data managers with OMOP CDM, expert interviews, and literature. Another strength was the use of the framework by Cabana et al [23]; although the categorization of the barriers was originally designed to review physician guideline adherence, we successfully applied this to adherence to FAIR principles, a specific type of guideline, since the framework was also designed to standardize the reporting of barriers to adherence. The classification gives insight into the nature of the barriers, highlighting areas that could be improved within a FAIRification process.

A limitation of this study is that a bias may be present in the scoring of the barriers by the data managers. The data managers may have been inclined to give their self-experienced barriers a higher score. However, as the top 10 contain a mix of the 3 sources, expected bias seems limited. Bias could further have been introduced as the data managers had to score each barrier even if they had little to no knowledge regarding that barrier. This limitation is as much as possible mitigated by providing the data managers more explanation when necessary during the first round and the live meeting of the Delphi method. Another limitation is that we did not focus on facilitators in this study. However, the recommendations we provide make the FAIR data implementation process easier and could therefore be considered as facilitators. Finally, it could be argued that a small subset of data

elements was used during the OMOP CDM implementation process, especially for the LDJ research site since they only had the PSQ as being mental health specific available in the chosen subset. However, we believe that the chosen subset was representative of the whole dataset to find all barriers. Even with introducing more data elements, it is likely that we would observe a repetition of already identified barriers instead of identifying new barriers. Moreover, not all barriers were related to data, some were technical, and focusing too much on the data aspect might have led to missing those technical barriers.

Future work should focus on designing a streamlined FAIR data process framework. This framework should contain practical and clear steps but should leave practical choices open to researchers or help them make informed decisions. Decision trees such as the one created by Verburg et al [36] could help researchers make their data FAIR. Their decision tree focuses on qualitative data; decision trees focused on other data or aspects of FAIR would be useful. The FAIR cookbook is a useful resource in this regard as well [42]. However, a FAIR recipe focused on making qualitative data FAIR is currently missing. The FAIR cookbook is written for expert users as it assumes familiarity with the FAIR principles, making it less of a starting point for beginners. Various communities have developed FAIR Implementation Profiles to provide guidance for implementing each principle [35,47,48], but a FAIR Implementation Profile focused on mental health data has not yet been developed, which would be relevant as further work.

In conclusion, 45 barriers that hamper the implementation of the FAIR principles were identified. The mental health domain does suffer largely from the same barriers to implementing the FAIR principles as other health care domains. We provided the top 10 barriers and recommendations for researchers to help them overcome these barriers. By providing this list, we hope that researchers are better prepared and motivated to start their FAIRification process and can better communicate their needs within their organization and towards grant providers. We strongly recommend the structural use of FAIR data stewards, creating clear step-bystep guides to make data FAIR, and the implementation of sustainable funding systems to maintain FAIR data infrastructure.

Acknowledgments

This research was funded by ZonMw (project number 50-56300-98-973). We also thank the COVID-19 Wellbeing and Mental Health Consortium for their valuable contribution to this study.

Conflicts of Interest

None declared.

Multimedia Appendix 1

Justification of using the Delphi method according to the Guidance on Conducting and REporting DElphi Studies (CREDES) by Jünger et al [24].

[DOCX File (Microsoft Word File), 18 KB-Multimedia Appendix 1]

Multimedia Appendix 2

All identified barriers and their type, affected FAIR principles, category, being specific to mental health, source, and total score. A forward slash in the source column indicates that the barrier was identified in multiple sources. [XLSX File (Microsoft Excel File), 17 KB-Multimedia Appendix 2]

Multimedia Appendix 3

All notes, scores, and comments of each Delphi method round. [XLSX File (Microsoft Excel File), 21 KB-Multimedia Appendix 3]

References

- Wilkinson MD, Dumontier M, Aalbersberg IJJ, et al. The FAIR Guiding Principles for scientific data management and stewardship. Sci Data. Mar 15, 2016;3(1):160018. [doi: 10.1038/sdata.2016.18] [Medline: 26978244]
- 2. Fair principles. URL: https://www.go-fair.org/fairprinciples/ [Accessed 2024-08-23]
- 3. van Reisen M, Stokmans M, Basajja M, Ong'ayo AO, Kirkpatrick C, Mons B. Towards the tipping point for FAIR implementation. Data Intelligence. Jan 2020;2(1-2):264-275. [doi: 10.1162/dint_a_00049]
- Kersloot MG, Abu-Hanna A, Cornet R, Arts DL. Perceptions and behavior of clinical researchers and research support staff regarding data FAIRification. Sci Data. May 27, 2022;9(1):241. [doi: <u>10.1038/s41597-022-01325-2</u>] [Medline: <u>35624282</u>]
- Garcia L, Bolleman J, Gehant S, Redaschi N, Martin M, UniProt Consortium. FAIR adoption, assessment and challenges at UniProt. Sci Data. Sep 20, 2019;6(1):175. [doi: <u>10.1038/s41597-019-0180-9</u>] [Medline: <u>31541106</u>]
- 6. Queralt-Rosinach N, Kaliyaperumal R, Bernabé CH, et al. Applying the FAIR principles to data in a hospital: challenges and opportunities in a pandemic. J Biomed Semantics. Dec 2022;13(1):1-19. [doi: 10.1186/s13326-022-00263-7]

- Lozano R, Fullman N, Mumford JE, et al. Measuring universal health coverage based on an index of effective coverage of health services in 204 countries and territories, 1990–2019: a systematic analysis for the Global Burden of Disease Study 2019. Lancet. Oct 2020;396(10258):1250-1284. [doi: 10.1016/S0140-6736(20)30750-9]
- Bacellar IOL, Morin G, Daniels S, Turecki G, Palaniyappan L, Lepage M. Opening up mental health research. J Psychiatry Neurosci. 2023;48(3):E209-E216. [doi: <u>10.1503/jpn.220199</u>] [Medline: <u>37253483</u>]
- Sadeh Y, Denejkina A, Karyotaki E, Lenferink LIM, Kassam-Adams N. Opportunities for improving data sharing and FAIR data practices to advance global mental health–Corrigendum. Glob Ment Health (Camb). 2023;10:e14. [doi: <u>10.</u> <u>1017/gmh.2023.14</u>]
- Boulkedid R, Abdoul H, Loustau M, Sibony O, Alberti C. Using and reporting the Delphi method for selecting healthcare quality indicators: a systematic review. PLoS One. 2011;6(6):e20476. [doi: <u>10.1371/journal.pone.0020476</u>] [Medline: <u>21694759</u>]
- 11. OHDSI. Standardized data: the OMOP common data model. URL: <u>https://www.ohdsi.org/data-standardization</u> [Accessed 2022-06-16]
- 12. Haverman L, van Oers HA, Limperg PF, et al. Implementation of electronic patient reported outcomes in pediatric daily clinical practice: the KLIK experience. Clin Pract Pediatr Psychol. 2014;2(1):50-67. [doi: 10.1037/cpp0000043]
- 13. DE LERENDE DATABANK JEUGD, DE-LERENDE-DATABANK-JEUGD.pdf [The learning database youth]. 2020. URL: <u>https://www.praktikon.nl/uploaded/publicaties/ De-Lerende-Databank-Jeugd.pdf</u>
- 14. Sejn. URL: <u>https://www.sejn.nl</u> [Accessed 2022-08-30]
- Zijlmans J, Tieskens JM, van Oers HA, et al. The effects of COVID-19 on child mental health: biannual assessments up to April 2022 in a clinical and two general population samples. JCPP Adv. Jun 2023;3(2):e12150. [doi: <u>10.1002/jcv2.</u> <u>12150</u>] [Medline: <u>37753155</u>]
- Fischer K, Tieskens JM, Luijten MAJ, et al. Internalizing problems before and during the COVID-19 pandemic in independent samples of Dutch children and adolescents with and without pre-existing mental health problems. Eur Child Adolesc Psychiatry. Oct 2023;32(10):1873-1883. [doi: 10.1007/s00787-022-01991-y]
- 17. Cantril H. Pattern of Human Concerns. Rutgers University Press; 1965.
- Vermulst A, Kroes G, De Meyer R, Nguyen L, Veerman JW. Handleiding OBVL, Eburon Uitgeverij BV [Manual PSQ (Parenting Stress Questionnaire)]. 2015.
- Klaufus LH, Luijten MAJ, Verlinden E, et al. Psychometric properties of the Dutch-Flemish PROMIS[®] pediatric item banks anxiety and depressive symptoms in a general population. Qual Life Res. Sep 2021;30(9):2683-2695. [doi: <u>10.</u> <u>1007/s11136-021-02852-y</u>] [Medline: <u>33983618</u>]
- 20. Extract transform load. URL: <u>https://ohdsi.github.io/TheBookOfOhdsi/ExtractTransformLoad.html</u> [Accessed 2023-07-19]
- 21. Data quality. URL: <u>https://ohdsi.github.io/TheBookOfOhdsi/DataQuality.html</u> [Accessed 2023-07-19]
- 22. OHDSI analytics tools. URL: <u>https://ohdsi.github.io/TheBookOfOhdsi/OhdsiAnalyticsTools.html</u> [Accessed 2023-07-19]
- 23. Cabana MD, Rand CS, Powe NR, et al. Why don't physicians follow clinical practice guidelines? A framework for improvement. JAMA. Oct 20, 1999;282(15):1458-1465. [doi: <u>10.1001/jama.282.15.1458</u>] [Medline: <u>10535437</u>]
- 24. Jünger S, Payne SA, Brine J, Radbruch L, Brearley SG. Guidance on Conducting and REporting DElphi Studies (CREDES) in palliative care: recommendations based on a methodological systematic review. Palliat Med. Sep 2017;31(8):684-706. [doi: 10.1177/0269216317690685]
- 25. Krippendorff K. Computing krippendorff's alpha-reliability. 2011. URL: <u>https://citeseerx.ist.psu.edu/document?repid=rep1&type=pdf&doi=de8e2c7b7992028cf035f8d907635de871ed627d</u> [Accessed 2024-03-21]
- 26. Krippendorff K. Content Analysis: An Introduction to Its Methodology. Sage Publications; 2018. [doi: 10.4135/ 9781071878781]
- 27. Wel of niet WMO-plichtig? Amsterdam Universitair Medische Centra. URL: <u>https://metc.amsterdamumc.org/wel-of-niet-wmo-plichtig/</u> [Accessed 2024-12-11]
- 28. Groenen KHJ, Jacobsen A, Kersloot MG, et al. The de novo FAIRification process of a registry for vascular anomalies. Orphanet J Rare Dis. Sep 4, 2021;16(1):376. [doi: 10.1186/s13023-021-02004-y] [Medline: 34481493]
- Overmars LM, Niemantsverdriet MSA, Groenhof TKJ, et al. A wolf in sheep's clothing: reuse of routinely obtained laboratory data in research. J Med Internet Res. Nov 18, 2022;24(11):e40516. [doi: <u>10.2196/40516</u>] [Medline: <u>36399373</u>]
- 30. Wise J, de Barron AG, Splendiani A, et al. Implementation and relevance of FAIR data principles in biopharmaceutical R&D. Drug Discov Today. Apr 2019;24(4):933-938. [doi: <u>10.1016/j.drudis.2019.01.008</u>] [Medline: <u>30690198</u>]

- Alharbi E, Gadiya Y, Henderson D, et al. Selection of data sets for FAIRification in drug discovery and development: which, why, and how? Drug Discov Today. Aug 2022;27(8):2080-2085. [doi: <u>10.1016/j.drudis.2022.05.010</u>] [Medline: <u>35595012</u>]
- 32. Alharbi E, Skeva R, Juty N, Jay C, Goble C. Exploring the current practices, costs and benefits of FAIR implementation in pharmaceutical research and development: a qualitative interview study. Data Intell. Oct 25, 2021;3(4):507-527. [doi: 10.1162/dint_a_00109]
- van Vlijmen H, Mons A, Waalkens A, et al. The need of industry to Go FAIR. Data Intell. Jan 2020;2(1-2):276-284. [doi: 10.1162/dint a 00050]
- 34. Arefolov A, Adam L, Brown S, et al. Implementation of the FAIR data principles for exploratory biomarker data from clinical trials. Data Intell. Oct 25, 2021;3(4):631-662. [doi: 10.1162/dint_a_00106]
- 35. Jacobsen A, de Miranda Azevedo R, Juty N, et al. FAIR principles: interpretations and implementation considerations. Data Intell. Jan 2020;2(1-2):10-29. [doi: 10.1162/dint_r_00024]
- 36. Verburg M, Braukmann R, Mahabier W. Making Qualitative Data Reusable— A Short Guidebook for Researchers and Data Stewards Working With Qualitative Data. 2023. [doi: <u>10.5281/zenodo.7777519</u>]
- Scholtens S, Jetten M, Böhmer J, et al. Final report: towards FAIR data steward as profession for the lifesciences. Report of a ZonMw funded collaborative approach built on existing expertise. 2022. URL: <u>https://doi.org/10.5281/zenodo.</u> 7225070 [Accessed 2023-07-19] [doi: <u>10.5281/zenodo.7225070</u>]
- Wendelborn C, Anger M, Schickhardt C. What is data stewardship? Towards a comprehensive understanding. J Biomed Inform. Apr 2023;140:104337. [doi: <u>10.1016/j.jbi.2023.104337</u>] [Medline: <u>36935012</u>]
- Engelhardt C, Biernacka K, Coffey A, et al. D7.4 how to be FAIR with your data. A teaching and training handbook for higher education institutions. Zenodo; 2022. URL: <u>https://doi.org/10.5281/zenodo.6674301</u> [Accessed 2023-07-24] [doi: <u>10.17875/gup2022-1915</u>]
- 40. Jacobsen A, Kaliyaperumal R, da Silva Santos LOB, et al. A generic workflow for the data FAIRification process. Data Intell. Jan 2020;2(1-2):56-65. [doi: 10.1162/dint_a_00028]
- 41. FAIRification Workflow. URL: <u>https://fairtoolkit.pistoiaalliance.org/methods/fairification-workflow</u> [Accessed 2024-08-28]
- 42. The FAIR cookbook. URL: https://faircookbook.elixir-europe.org/content/home.html [Accessed 2023-07-24]
- 43. Rocca-Serra P, Sansone SA. Experiment design driven FAIRification of omics data matrices, an exemplar. Sci Data. Dec 12, 2019;6(1):271. [doi: 10.1038/s41597-019-0286-0] [Medline: 31831744]
- 44. USCDI+. URL: https://www.healthit.gov/topic/interoperability/uscdi-plus [Accessed 2024-03-21]
- 45. van Damme P, Moreno PA, Bernabé CH, et al. A resource for guiding data stewards to make European rare disease patient registries FAIR. Data Sci J. Jun 5, 2023;22. [doi: 10.5334/dsj-2023-012]
- 46. Mayer G, Müller W, Schork K, et al. Implementing FAIR data management within the German Network for Bioinformatics Infrastructure (de.NBI) exemplified by selected use cases. Brief Bioinform. Sep 2, 2021;22(5):bbab010. [doi: <u>10.1093/bib/bbab010</u>] [Medline: <u>33589928</u>]
- 47. FAIR. FAIR implementation profile. URL: <u>https://www.go-fair.org/how-to-go-fair/fair-implementation-profile/</u> [Accessed 2023-07-19]
- Schultes E, Magagna B, Hettne KM, Pergl R, Such´anek M, Kuhn T. Reusable FAIR implementation profiles as accelerators of FAIR convergence. Presented at: International Conference on Conceptual Modeling; 2020:138-147; [doi: 10.1007/978-3-030-65847-2_13]

Abbreviations

CDM: Common Data Model CREDES: Guidance on Conducting and REporting DElphi Studies DREAMS: Dutch REsearch in child and Adolescent Mental health DSM: Diagnostic and Statistical Manual of Mental Disorders DSM-5: Diagnostic and Statistical Manual of Mental Disorders (Fifth Edition) ETL: Extract, Transform, Load FAIR: Findable, Accessible, Interoperable, Reusable GDPR: General Data Protection Regulation LDJ: Dutch abbreviation for Learning Database Youth OHDSI: Observational Health Data Sciences and Informatics OMOP: Observational Medical Outcomes Partnership PROMIS: patient-reported outcomes measurement information system PSQ: Parenting Stress Questionnaire

USCDI+: United States Core Data for Interoperability Plus

Edited by John Torous; peer-reviewed by Ashley Knapp, Mark Zmavc, Oscar Tamburis; submitted 02.04.2024; final revised version received 23.09.2024; accepted 08.10.2024; published 19.12.2024

Please cite as:

de Groot R, van der Graaff F, van der Doelen D, Luijten M, De Meyer R, Alrouh H, van Oers H, Tieskens J, Zijlmans J, Bartels M, Popma A, de Keizer N, Cornet R, Polderman TJC Implementing Findable, Accessible, Interoperable, Reusable (FAIR) Principles in Child and Adolescent Mental Health Research: Mixed Methods Approach JMIR Ment Health 2024;11:e59113 URL: <u>https://mental.jmir.org/2024/1/e59113</u> doi: <u>10.2196/59113</u>

© Rowdy de Groot, Frank van der Graaff, Daniël van der Doelen, Michiel Luijten, Ronald De Meyer, Hekmat Alrouh, Hedy van Oers, Jacintha Tieskens, Josjan Zijlmans, Meike Bartels, Arne Popma, Nicolette de Keizer, Ronald Cornet, Tinca J C Polderman. Originally published in JMIR Mental Health (<u>https://mental.jmir.org</u>), 19.12.2024. This is an open-access article distributed under the terms of the Creative Commons Attribution License (<u>https://creativecommons.org/licenses/by/4.0/</u>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Mental Health, is properly cited. The complete bibliographic information, a link to the original publication on <u>https://mental.jmir.org/</u>, as well as this copyright and license information must be included.