Sharing brain imaging data in the Open Science era: how and why?

Kathrin Giehl, Henk-Jan Mutsaerts, Kristien Aarts, Frederik Barkhof, Svenja Caspers, Gaël Chetelat, Marie-Elisabeth Colin, Emrah Düzel, Giovanni B Frisoni, M Arfan Ikram, Jorge Jovicich, Silvia Morbelli, Wolfgang Oertel, Christian Paret, Daniela Perani, Petra Ritter, Bàrbara Segura, Laura E M Wisse, Elke De Witte, Stefano F Cappa, Thilo van Eimeren

The sharing of human neuroimaging data has great potential to accelerate the development of imaging biomarkers in neurological and psychiatric disorders; however, major obstacles remain in terms of how and why to share data in the Open Science context. In this Health Policy by the European Cluster for Imaging Biomarkers, we outline the current main opportunities and challenges based on the results of an online survey disseminated among senior scientists in the field. Although the scientific community fully recognises the importance of data sharing, technical, legal, and motivational aspects often prevent active adoption. Therefore, we provide practical advice on how to overcome the technical barriers. We also call for a harmonised application of the General Data Protection Regulation across EU countries. Finally, we suggest the development of a system that makes data count by recognising the generation and sharing of data as a highly valuable contribution to the community.

Lancet Digit Health 2024; 6: e526–35

Department of Nuclear Medicine (K Giehl PhD, Prof T van Eimeren MD), **and Department of Neurology** (Prof T van Eimeren), **Faculty of Medicine and University Hospital Cologne, University of Cologne, Cologne, Germany; Institute of Neurosciences and Medicine (INM-2)** (K Giehl), **and Institute of Neuroscience and Medicine (INM-1)** (Prof S Caspers PhD), **Research Center Jülich, Jülich, Germany; Radiology and Nuclear Medicine, Vrije Universiteit Medical Center, Amsterdam University Medical Center, Amsterdam, Netherlands** (H-J Mutsaerts PhD, Prof F Barkhof MD); **Amsterdam Neuroscience, Brain Imaging, Amsterdam, Netherlands** (H-J Mutsaerts, Prof F Barkhof); **European Brain Council, Brussels, Belgium** (K Aarts PhD, Prof W Oertel MD); **Queen Square Institute of Neurology and Centre for Medical Image Computing, University College London, London, UK** (Prof F Barkhof); **Institute for Anatomy I, Medical Faculty and University Hospital Düsseldorf, Heinrich Heine University Düsseldorf, Düsseldorf, Germany** (Prof S Caspers); **Normandie université, UNICAEN, INSERM, U1237, NeuroPresage Team, Cyceron, Caen, France** (G Chetelat PhD); **EBRAINS, Brussels, Belgium** (M-E Colin MSc); **Faculty of Medicine, Institute for Cognitive Neurology and Dementia Research, University of Magdeburg, Magdeburg, Germany** (Prof E Düzel MD); **Center for Neurodegenerative Diseases, Magdeburg, Germany** (Prof E Düzel); **Department of Rehabilitation and Geriatrics, Memory Center, Geneva University and University Hospitals, Geneva, Switzerland** (Prof G B Frisoni MD); **Department of Epidemiology** (Prof M A Ikram PhD), **and Neurosurgical Department**

Introduction

The identification of reliable brain imaging biomarkers is pivotal for improving our understanding of neurological and psychiatric disorders, earlier detection of disease, and the monitoring of potentially disease-modifying treatments. Data sharing between individual scientists and institutes regionally, nationally, within Europe, and globally will probably accelerate the development of suitable imaging biomarkers and thus foster the translation of recent scientific advances into clinical practice, ultimately improving patient care.

The sharing of research data is imperative under the European Code of Conduct for Research Integrity.[1] Data sharing enhances reproducibility of studies via increased transparency of procedures and processes, allows for more variety of analyses of existing datasets, and increases the quantity and diversity of data for artificial intelligence approaches and big data science. Altogether, sharing research data increases the quality of research overall.[2] Contributing to high-quality research is a prime motivation for the scientific community; however, the benefits of sharing data beyond intrinsic motivation might be less obvious to individual scientists. Furthermore, research is traditionally more single-centre oriented, so it might not be clear how research groups and institutions can overcome traditional sharing rules and contribute data in an open science context. In other words, how and why should we share?[3]

To address this unmet need, the European Cluster for Imaging Biomarkers (ECIB) was established within the EU-funded European Brain Research Area in May, 2021. The ECIB pursues two major objectives: to provide a systematic examination of the viewpoints, positions, and approaches with regard to opportunities and challenges of sharing brain imaging data, particularly considering the requirements of EU data protection laws, and to give recommendations on sharing brain imaging data to the brain imaging community, policy makers, and funding agencies on local, national, and European levels.

Although the principles discussed in this position statement might be applicable to other types of data as well, this Health Policy primarily focuses on brain imaging data—or data in brain space—that can be acquired using modalities such as MRI, CT, and PET. The recommendations provided here follow the consensus reached by the members of ECIB (appendix p 2) on the basis of the results of a survey conducted among principal investigators of large European neuroimaging consortia, thereby representing European brain imaging experts. All ECIB members asked all principal investigators of their respective consortia to participate in the online survey (appendix p 1).

In this Health Policy, we summarise the main survey results and discuss the identified main sharing challenges in detail—technical data sharing aspects, the complexity of legal barriers in the context of the EU data protection law, and the scarcity of data sharing incentives. Finally, we provide recommendations to scientists and necessary requests to policy makers to improve data sharing processes. A glossary of terms is provided in the panel.

Summary of survey results

The survey was compiled by members of the ECIB initiative on the basis of previous work,[6] implemented using SurveyMonkey, and open between Dec 21, 2021, and March 3, 2022. Here, we summarise the most striking findings (for more details see appendix pp 1–2).

Participant demographics

95 people responded to the survey overall. Of the 72 survey questions, not all were mandatory, resulting in differing numbers of responses for each survey question. The average respondent was male (which is reflective of the sex distribution of male and female professors across the EU[7]), was aged 38–58 years, and had 12–28 years scientific experience. 42 (69%) of the 61 survey respondents who provided demographic information

(E De Witte PhD), **Erasmus University Medical Center, Rotterdam, Netherlands; Center for Mind/Brain Sciences, University of Trento, Rovereto, Italy** (Prof J Jovicich PhD); **Nuclear Medicine Unit, AOU Città Della Salute e Della Scienza di Torino, Turin, Italy** (Prof S Morbelli PhD); **Department of Medical Sciences, University of Turin, Turin, Italy** (Prof S Morbelli); **Department of Neurology, University of Marburg, Marburg, Germany** (Prof W Oertel); **Department of Psychosomatic Medicine and Psychotherapy, Central Institute of Mental Health Mannheim, Medical Faculty Mannheim/Heidelberg University, Mannheim, Germany** (C Paret PhD); **San Raffaele University and San Raffaele Scientific Institute, Milan, Italy** (Prof D Perani MD); **Berlin Institute of Health** (Prof P Ritter PhD), **and Department of Neurology with Experimental Neurology** (Prof P Ritter), **Charité, Universitätsmedizin Berlin, Berlin, Germany; Bernstein Focus State Dependencies of Learning and Bernstein Center for Computational Neuroscience, Berlin, Germany** (Prof P Ritter); **Einstein Center for Neuroscience Berlin, Berlin, Germany** (Prof P Ritter); **Einstein Center Digital Future, Berlin, Germany** (Prof P Ritter); **Medical Psychology Unit, Department of Medicine, Institute of Neurosciences, University of Barcelona, Barcelona, Spain** (B Segura PhD); **Hospital Clinic Foundation for Biomedical Research–August Pi i Sunyer Biomedical Research Institute, Barcelona, Spain** (B Segura); **Biomedical Research Networking Center on Neurodegenerative Diseases Barcelona, Spain** (B Segura); **Department of Clinical Sciences Lund, Lund University, Lund, Sweden** (L E M Wisse PhD); **University Institute of Advanced Studies, Pavia, Italy** (Prof S F Cappa MD); **IRCCS Mondino Foundation, Pavia, Italy** (Prof S F Cappa)

***Panel:* Glossary**

Brain imaging data structure (BIDS)
In 2016, the implementation of a standardised BIDS was proposed to systematically organise data from neuroimaging experiments.[4] The aim of BIDS is to increase the transparency of the data structure, making it easier to share data with other scientists and to make data interoperable for data software analyses packages.

Data transfer agreement (DTA) or data use (or user) agreement
Although often used synonymously, a data transfer agreement refers to the transfer of data between parties, and a data user agreement regulates the permissible uses of the transferred data by the recipient.

The General Data Protection Regulation (GDPR)
The GDPR is applicable law in the European Union as of 2018 and protects people with regard to the processing of personal data.

Findable, accessible, interoperable, and reusable (FAIR)
The FAIR guidelines were introduced in 2016 and aim to optimise and support the reuse of scientific data, and to produce findable, accessible, interoperable, and reusable data. FAIRness can be achieved by making metadata findable and giving ways to make data accessible under a DTA or data user agreement; if the data is organised in BIDS format, it becomes interoperable and reusable. Therefore, in principle, data can be considered FAIR without the raw data necessarily being open.

Preregistration
The practice of making a research plan, including hypotheses, methods, and analysis approaches, public.

Open access
The practice of providing online access to scientific information that is free of charge and reusable to the user. This information includes peer-reviewed publications, data underlying publications, and other datasets.[5]

Open science
An approach to research based on open cooperative work that emphasises the sharing of knowledge, results, and tools as early and as widely as possible.[5]

were male and 44 (72%) were from the EU. Of these 61 participants, 49 (80%) had attained assistant professor level or higher, 39 (64%) were associated with a medical faculty, and 33 (54%) were trained as medical doctors. Their primary research fields were neurology (27 [44%]) or cognitive science (13 [21%]).

The 77 respondents who answered questions about their study methods mainly studied patients with neurodegenerative disorders (64 [83%]), at-risk populations (40 [52%]), and healthy individuals (51 [66%]) in cross-sectional (25 [32%]) or longitudinal (52 [68%]) study designs. Imaging methods used by respondents included MRI, PET, and single-photon emission computed tomography (SPECT). 56 (73%) of these 77 respondents used anatomical MRI; 51 (66%) used diffusion MRI, 47 (61%) used resting state functional MRI (fMRI), 27 (35%) used task-based fMRI, 35 (45%) used amyloid PET, 19 (25%) used tau PET, 33 (43%) used fluorodeoxyglucose tracers, and 22 (29%) used SPECT.

Most of the 55 respondents who provided funding information were involved in one or more major European funding programmes—38 (69%) were involved in EU Framework Programmes, 36 (65%) in the EU Joint Programme – Neurodegenerative Disease Research, 21 (38%) in the Innovative Medicines Initiative, ten (18%) in the Human Brain Project, and five (9%) in the Network of European Funding for Neuroscience Research.

Preregistration

68 respondents answered the questions about preregistration. 46 (68%) of these 68 respondents agreed that preregistration increases the credibility of data, 56 (82%) agreed it increases the visibility of ongoing projects, and 51 (75%) agreed it enhances the quality of research (figure 1); however, only 23 (33%) of the respondents supported making preregistration mandatory. This reservation might be because, despite the acknowledgment of the potential preregistration benefits, 30 (44%) of 68 respondents had never preregistered a project before. There were several main reasons against preregistration. 46 (68%) of 68 respondents stated they had insufficient time to preregister, and 34 (50%) stated they had insufficient knowledge of how to adequately preregister. 37 (54%) of 68 respondents had insufficient knowledge of suitable preregistration platforms, and 37 (54%) had insufficient knowledge of how preregistration might affect publication and data sharing in a highly competitive scientific environment.

Brain imaging data structure

68 respondents answered the questions about brain imaging data structure (BIDS). 46 (68%) of these 68 survey respondents had heard about BIDS. Of the 46 respondents familiar with BIDS, 28 (61%) used it. 40 respondents answered the multiple-choice question on possible barriers regarding BIDS implementation. Answers included having insufficient time (eight [20%] of 40 respondents), technical knowledge (six [15%] of 40 respondents), or human power (four [10%] of 40 respondents) to convert data into BIDS. 54 (81%) of the 67 respondents who provided additional answers to questions relating to BIDS indicated their willingness to use BIDS once their preferred analysis software was

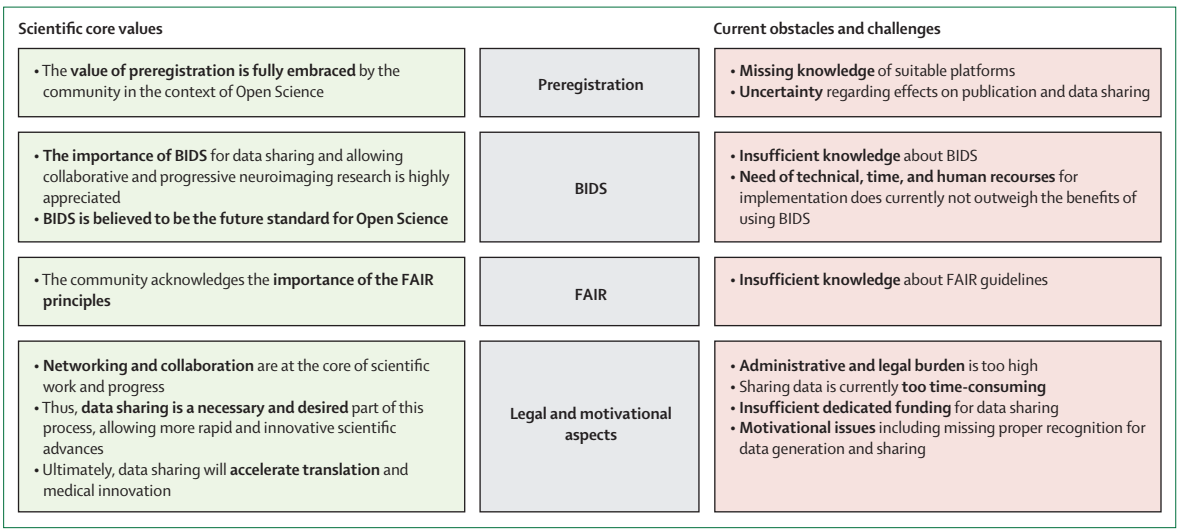

Figure 1: **Values of the neuroimaging community and obstacles and challenges in the context of Open Science**
BIDS=brain imaging data structure. FAIR=findable, accessible, interoperable, reuseable.

Correspondence to:
Prof Thilo van Eimeren, Department of Nuclear Medicine, Faculty of Medicine and University Hospital Cologne, University of Cologne, Cologne 50937, Germany
thilo.van-eimeren@uk-koeln.de

For more on the **European Cluster for Imaging Biomarkers** see https://www.ebra.eu/ecib-cluster/

See **Online** for appendix

For more on **Brain Imaging Data Structure** see https://bids.neuroimaging.io

BIDS-compatible. Four (6%) of these 67 respondents were not convinced that BIDS will become the standard in neuroimaging, and two (3%) dismissed the future use of BIDS in their work environment.

Findable, accessible, interoperable, and reuseable

45 (70%) of the 64 respondents who answered questions about the principles of findability, accessibility, interoperability, and reuse of digital assets were aware of these principles, but only 30 (47%) had implemented the principles in their projects. 58 (91%) of these 64 respondents recognised the importance of their data being findable, accessible, interoperable, and reusable (FAIR). 57 (89%) of these respondents envisioned adhering to the FAIR principles in the future.

Experience with data sharing

61 respondents answered questions related to their experience with data sharing. 59 (97%) of these 61 respondents indicated they had shared data with scientists outside their institution. In addition, 57 (93%) agreed that data sharing fosters communication and cooperation between research groups, 59 (97%) agreed it enhances the generalisability of results, and 59 (97%) agreed it increases statistical power for providing additional or more reliable conclusions. 59 (97%) agreed that data sharing increases the usability of their data and 51 (84%) of the 61 respondents found multicentre studies more reliable than monocentre studies.

Although 53 (86%) of the 61 respondents disagreed with the statement that they prefer not giving other researchers access to their raw data, the inclination to share data depended on the level of processing (ie, metadata or data derivates were more likely to be shared than source data). Managed repositories with restricted access were preferred by 50 (82%) of the 61 respondents compared with completely open repositories (19 [31%]). 42 (69%) of the 61 respondents seemed to be willing to share data upon personal request, and 48 (79%) preferred data to be shared under a data transfer agreement (DTA).

Experience with legal barriers

61 respondents answered questions regarding their experience with legal barriers. The administrative burden and time needed to set up a DTA that complies with the European General Data Protection Regulation (GDPR) was regarded as too high by 42 (69%) of these 61 respondents. 27 (44%) felt that the current legal constraints prevent them from sharing data, either because of experiencing or fearing Institutional Review Board rejection (22 [36%]) or because anonymity could not be fully achieved (16 [26%]). 22 (36%) indicated that they were unlikely to share their research data given the current legal constraints, but 61 (100%) of these 61 survey respondents indicated their willingness to share research data if there were no legal or technical constraints.

Motivation to share neuroimaging data

In addition to the heavy administrative burden, sharing data seemed to be associated with obstacles, fears, and barriers for the 61 survey respondents who answered questions relating to sharing neuroimaging data. Although these barriers were not related to the know-how of sharing or the data's complexity or vast size, data sharing was perceived as a very costly act in terms of time and human resources by 43 (70%) of the 61 respondents, and 49 (80%) indicated that there is a lack of funding to make data suitable for online sharing.

Although only six (10%) of the 61 respondents were afraid that others could detect errors in their data and

11 (18%) were afraid that others could perform alternative analyses refuting their initial conclusions, 25 (41%) feared that others could publish results on their data before they could, and 24 (39%) feared they would not get proper recognition for data sharing. If reasonable compensation (ie, adequate scientific recognition) was provided, 42 (69%) of the 61 respondents would be happy to share.

Discussion

In the following section, the technical, legal, and motivational challenges are addressed in greater detail and recommendations and requests to policy makers and funding agencies to overcome these in the future are provided (figure 2).

Technical aspects

Preregistration

For more on **Open Science Framework preregistration** see https://osf.io/

For more on **AsPredicted** see https://aspredicted.org/

For more on **Bio-protocol** see https://bio-protocol.org/en

For more on the **ISRCTN registry** see https://www.isrctn.com/

Preregistration involves the publication of the study protocol before data collection in a repository or registry. Alternatively, Registered Reports involve submitting the theoretical background, study design, methods, and planned analysis as a type of journal article for publication. Registered Reports offer notable advantages. First, this kind of article will be peer-reviewed, therefore encouraging scientific discussion and potentially reducing possible results-based critiques from future peer reviewers. Second, once the study is accepted by the journal, the final results will be published regardless of the outcome, after verification by peer review that the research accomplished the registered research plan, as detailed in a preprint by Stewart and colleagues.[8]

Other advantages of preregistration and Registered Reports include the reduction of hypothesising after the results are known,[9] data dredging or p-hacking (ie, the misuse of data analysis approaches to find results that can be presented as statistically significant),[10] and publication bias.[11] Therefore, Registered Reports incentivise publishing non-statistically significant results in addition to positive, novel, and attractive outcomes. Preregistration—regardless of whether data are being shared—is an indispensable practice in the light of open science, helping to reduce competitiveness in the research system and thereby promoting cooperation among researchers and stakeholders.

Recommendations and resources for preregistration

The detailing of the study plan registry or publishing a full study protocol in an academic journal is an effort for the researcher. To help with this task, several online templates are available, tailored to specific disciplines and types of research designs. Examples can be found in the Open Science Framework (OSF) preregistration, on AsPredicted, in Bio-protocol for biological science studies, and in ISRCTN registry for clinical research. In the neuroimaging field, OSF provides a specific fMRI preregistration template.[12] Regarding Registered Reports, an important source of information is The Center for Open Science,[13] which provides a ten-item questionnaire to guide researchers in stage 1 manuscript development and maintains an up-to-date list of journals accepting Registered Reports with a variety of requirements for submission.

Even in the most planned and anticipated research, deviations from initial methods and analysis plans are common. Deviations do not necessarily rule out testing predictions effectively but should be reported (otherwise, preregistration of the study is worthless). Nevertheless, although preregistration documents prespecified research plans before seeing results, it should not be considered a restriction to research. Preregistered

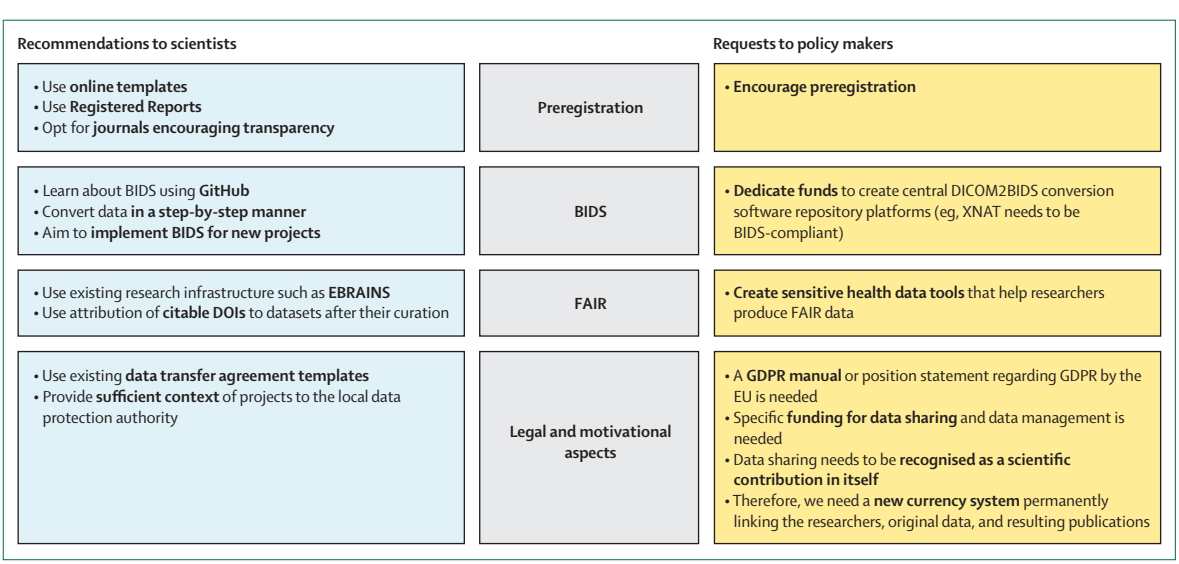

Figure 2: **Recommendations and requests for fostering the sharing of neuroimaging data in an Open Science framework**
BIDS=brain imaging data structure. DOIs=digital object identifier. FAIR=findable, accessible, interoperable, reuseable. GDPR=General Data Protection Regulation.

protocols can be updated before and during data collection, and before data analyses. Additionally, exploration of the data and testing of hypotheses that were not preregistered is possible, as long as such hypotheses and analyses are clearly labelled as a posteriori in the paper.[14]

Harmonisation

Harmonisation of neuroimaging practices is driven by the need to improve the reproducibility of scientific results derived from neuroimaging data and is a crucial step for successful data sharing. However, preclinical neuroimaging data acquisition and analyses are limited by the heterogeneity of neuroimaging hardware and software and the absence of standardised or harmonised acquisition, quality control, and analysis protocols. Accordingly, upfront standardisation of imaging parameters needs to be considered in prospective multicentre neuroimaging studies.

Ensuring consistency in neuroimaging acquisition protocols is vital but complex due to the evolving nature of neuroimaging technologies, which presents a constantly shifting landscape.[15] This challenge is compounded by variations not only across different vendors, but also within each vendor's own developments of hardware, pulse sequence programs, and image reconstruction programs.

Besides acquisition harmonisation issues, to increase the use of scientific data and encourage neuroimaging harmonisation efforts, increasing the sharing of well characterised derived metrics (with details of the pipelines used to derive them) that are meaningful and clinically relevant to the wider clinical research community is also crucial.

Recommendations and resources for harmonisation

Several international harmonisation initiatives have been conducted for both single modalities (eg, MRI, PET, or SPECT) and multimodal imaging pipelines. Although these initiatives have not solved intrinsic imaging-related variability issues, they have contributed to the definition of harmonisation goals.

Such initiatives include—but are not limited to—the Committee on Best Practices in Data Analysis and Sharing,[16] the Open Science Initiative for Perfusion Imaging,[17] the Quantitative Imaging Biomarkers Alliance,[18] Harmonizing Brain Imaging Methods for Vascular Contributions to Neurodegeneration,[19] and the Worldwide Alzheimer's Disease Neuroimaging Initiative.[4,20] The Committee on Best Practices in Data Analysis and Sharing provides best practices for image analysis, results reporting, and algorithm and data sharing on behalf of the Organization for Human Brain Mapping. The Open Science Initiative for Perfusion Imaging is an initiative of the International Society of MRI perfusion study group that focuses on the standardisation of acquisition and processing terminology and inventories for data and pipelines. The International Society of MRI Reproducibility group supports open-source and scanner-independent imaging frameworks[21] for the development and use of pulse sequences (eg, Pulseq[22] or QMRpullseq[23]) and image reconstruction tools (eg, Gadgetron[24]). The Radiological Society of North America's Quantitative Imaging Biomarkers Alliance aims to evaluate imaging biomarker performance for clinical trials. Harmonizing Brain Imaging Methods for Vascular Contributions to Neurodegeneration is an initiative funded by the EU Joint Programme – Neurodegenerative Disease Research that provides a website with harmonised acquisition protocols, a software database, rating scales, and case report forms. The goal of the Worldwide Alzheimer's Disease Neuroimaging Initiative is to harmonise projects and results across different geographical regions by harmonising data management and availability.

BIDS

BIDS provides a systematic structure for organising neuroimaging metadata in a fashion that is both intuitive for humans and optimised for machine readability, as detailed in a preprint by Poldrack and colleagues.[25] BIDS was initially developed for MRI data, but has been extended to accommodate a wide range of imaging modalities, including electroencephalogram,[26] magnetoencephalogram,[27] PET,[28] and countless others. Many BIDS extension proposals are in development (eg, for genetic and other clinical non-imaging data).

A wide range of neuroimaging analysis software has already been made BIDS compatible, including popular packages such as Freesurfer, FMRIB Software Library, Statistical Parametrical Mapping, and Analysis of Functional NeuroImaging, available in defined software packages (ie, Dockers) as so-called BIDS apps.[29]

BIDS recognises three types of data. First, BIDS recognises the source data containing the data as they come from the scanners. Second, BIDS recognises the raw data containing the original data in BIDS structure, where each image in Neuroimaging Informatics Technology Initiative format and JSON metadata sidecar files have been curated, validated, and quality controlled. These data are shared, ensuring that everyone uses the same raw data. Third, BIDS recognises the derivatives containing pipeline outputs, which can range from images such as tissue segmentations to tabular data such as tissue volume values. BIDS derivatives have a similar structure to the raw data, with the addition of provenance, a list of processing steps that were undertaken to produce the derivatives.

Although BIDS itself is a standard, the Digital Imaging and Communications in Medicine (DICOM) images coming from scanners are far from standardised. The preparation of BIDS data can therefore be a major challenge. Although most BIDS parameters can be found in the DICOM header, some metadata need to be

retrieved from general scanner protocols or from the exact sequence specifications on the scanner itself. In the future, BIDS will probably continue to evolve and expand to meet the needs of the medical and neuroimaging communities. One area of development is the integration of BIDS with other data standards and ontologies, such as the Neuroimaging Data Model[30] and the Committee on Best Practices in Data Analysis and Sharing.[16]

Recommendations and resources for BIDS

Although BIDS might be daunting to a first-time user, it is now widely used in the neuroimaging community. BIDS is open-source and has three main parts that are managed on Github: the BIDS specification itself,[31] the BIDS validator,[32] and BIDS examples.[33] This combination allows the user to learn the philosophy and reasoning behind BIDS, upload a dataset to test its BIDS compliance, and study example datasets. The BIDS validator can be especially useful to ascertain that no metadata are missing, even for future potential analyses. A growing number of journals require data to be shared in BIDS format, and many imaging centres and data repositories have adopted BIDS as their preferred format. Therefore, we recommend that investigators convert their data to BIDS using a step-by-step approach.

FAIR principles

Many institutions and initiatives around the world are setting up infrastructures for the storage of data allowing data reuse, thereby increasing the value of the data produced. To improve data sharing, the FAIR guiding principles were introduced in 2016;[34,35] however, barriers remain, as identified in the Towards European Health Data Space report published in February 2022.[36] One of these barriers is the use of different interoperability standards across Europe, which makes comparisons and sharing data and research results challenging. This report also shows that poor data management procedures reduce the capability to reuse the data.

Recommendations and resources for FAIR

An increasing number of funders are requiring data management plans, and journals require individual researchers to publish their data so that they can be inspected and reused. The Human Brain Project has created some resources to help facilitate meeting these requirements, which are now implemented in the digital research infrastructure EBRAINS. These resources include DataLad, the virtual research environment,[37] and openMINDS, a metadata framework that develops and maintains a set of metadata models, libraries of controlled terminologies, brain atlases, common coordinate spaces for neuroscience graph databases, and the attribution of citable DOIs to datasets after their curation.[38] Nevertheless, sensitive health data tools are currently unable to help researchers produce FAIR data. The implementation of such tools, however, is clearly needed with the creation of the European Data Health Space by the European Commission.

Legal aspects

The balance between data privacy risk management and open science initiatives has been widely recognised as a major challenge for neuroscience researchers.[39–41] Data privacy issues have a huge effect on neuroimaging research,[42] and the workload required to comply with the European GDPR[43] is considered an important obstacle to data sharing.

The GDPR is directly applicable by law without the need of transposition into federal law by EU member states, and thus represents a step toward the harmonisation of legislation regarding personal data protection within the EU.[44] Importantly, the GDPR also provides flexibility for individual member states to modify or derogate from some of its provisions. Thus, the balance between data privacy risk management and scientific potential is subjective, making the actual implementation of GDPR in informed consent forms and DTAs challenging for both researchers and legal representatives.

The main ethical principle behind the GDPR is to empower research participants to make informed decisions about the use and possible reuse of their data. A frequently reported sentiment, particularly among researchers involved in the secondary use of personal health data, is that the GDPR's complexity and rigidity result in a restriction of public data sharing and, thus, a restriction of scientific progress.[45,46] This sentiment is paradoxical, considering the clear statement in the GDPR that "the free movement of personal data within the Union shall be neither restricted nor prohibited for reasons related to personal data protection".[47] A typical restrictive approach to the reuse of data (eg, by the Dutch or Italian Data Protection Authority) is, in short, a strict interpretation of the written informed consent of a living human being, in which each research aim and each potential data-receiving party should be explicitly specified. Exceptions need to be evaluated by the relevant data protection authority (eg, whether tracing a large number of participants is impractical).[48] Again, the stance of the GDPR on this issue is clear: "further processing for archiving purposes in the public interest, scientific or historical research purposes or statistical purposes shall [...] not be considered to be incompatible with the initial purposes ('purpose limitation')".[47] A more liberal approach is the so-called secondary informed consent, in which research volunteers are also able to state in their informed consent form that they allow data sharing and reuse for future related research aims.

Anonymisation challenges

The GDPR only concerns data that are in some way identifiable (ie, personal data). Therefore, one solution for GDPR compliance would be to fully anonymise or

deidentify collected data before sharing; however, the ability to identify data and thus classify it as personal data to which the GDPR applies, or to have truly anonymous data, is a subject of debate. A clear case is the anonymous distribution of a dataset containing a single patient with a rare disease. Such a special case would be easy to identify. At the other end of the spectrum would be a large multisite dataset with healthy volunteers. Therefore, the ability to anonymously share data without privacy concerns (ie, without the applicability of GDPR) depends on the interpretation of the risk of identification. For GDPR, this risk is based on an estimation: the data are considered anonymised if the effort required for reidentification is estimated to be unreasonable. This definition contrasts with the definition of other regulatory regimes, such as the US Health Insurance Portability and Accountability Act, which specifies a list of 18 identifiers (eg, names, email addresses, and social security numbers) to be deleted in order for the data to be considered anonymous.[49] Although the GDPR estimation-based approach might be a source of confusion, the latter normative approach does not easily integrate technical developments, resulting in increased possibility of reidentification. For example, one challenge specific to neuroimaging data is that structural images containing facial information might facilitate patient identification, especially with recently developed deep learning methods;[50] however, the efficacy of these potential tools for reidentification has not yet been tested outside of academia. To date, there are no known incidents of threat actors having exploited brain imaging data, which shows the crucial need to assess the actual likelihood of reidentification under more realistic conditions, according to a preprint by Jwa and colleagues.[51] Although defacing methods exist to counter this issue,[52,53] including practices detailed in a preprint by Clunie and colleagues,[54] they vary in their effectiveness and might hamper subsequent image-processing pipelines.[55]

Recommendations and resources for anonymisation challenges

To expedite legal proceedings involving data sharing, understanding the ethics and basic rules behind data sharing requirements is important for researchers. Directly engaging with the legal representatives and data privacy officers of the institution can be helpful for researchers to understand their principles, motivation, and restrictions. The use of existing DTAs and informed consent templates should be encouraged. These templates should be shared with a researcher's legal departments and data privacy officers. The Open Brain Consent, for example, is an international initiative striving for standardisation in this context that has proposed a data user agreement template that allows data sharing and reuse while complying with GDPR.[56] Importantly, scientists should provide as much context to their legal officers as possible. This context will help legal officers to understand the intention of the data sharing, possibly expediting the creation of a suitable DTA. Likewise, a non-legal summary might be beneficial, for example, at the top of a DTA to increase understanding of partners and collaborators.

Overall, however, a GDPR manual, position statement, or explanation written by the EU, ideally also including information about the requirements for data sharing with countries under a different regulatory regime, would be of great benefit to the scientific community. Most scientists understand that risk management is subjective and find providing a work environment in which privacy is protected to be important, but find understanding the fundamentals of the GDPR to be challenging.

Motivational aspects

Scientists are motivated by a variety of factors, including intrinsic satisfaction in solving complex problems (sometimes referred to as puzzle), academic reputation (ribbon), and career-related benefits such as increased salary or entrepreneurial success (gold).[57] The latter two factors are closely tied to how a researcher's contributions to scientific progress are externally evaluated. High-impact factor journal publications have historically played a central role in this evaluation process within a traditional academic evaluation framework.

Comparable to the agricultural cycle, a research project follows a sequence of stages. Initial groundwork (comparable to preparing the soil in this agricultural analogy) includes previous work, grant applications, and administrative tasks encompassing legal, ethical, and operational considerations. Subsequently, resources are invested in building the infrastructure of the project, forming the research team, and recruiting participants (akin to sowing the seeds). The ongoing project management, quality assurance, and data management mirror the continuous care needed in the growth phase (comparable to irrigation, fertilizing, and weeding). Finally, the data analysis phase (drawing parallels with harvesting) concludes the project, with the goal being the publication of results. Essentially, this academic harvest can translate into academic reputation and career advancements; however, with the open science paradigm, researchers are expected to share their data with the broader research community as soon as possible. This aspect creates a noteworthy discrepancy between the effort invested (fieldwork) and the resulting outcomes (harvest), which raises a crucial question: what is the incentive to generate data and make data universally accessible? Relying solely on the altruism of certain researchers might be a simplistic and imprudent stance. Consequently, requiring researchers to share data during project funding might have counterproductive outcomes for scientific progress—unless mechanisms are devised to incentivise data provision in a way that meaningfully advances academic careers.

Historically, incentives for data sharing have commonly been tied to authorship on publications. Noteworthy

examples, such as the Alzheimer's Disease Neuroimaging Initiative (ADNI) and the Parkinson's Progression Marker Initiative, have implemented this approach. ADNI, a multisite study funded through a public–private partnership, has made data available without any embargo period since 2004, resulting in more than 2500 publications. The ADNI group of investigators is acknowledged as a coauthor in each of these works. This approach faces several challenges, however. The concept of these group authorships can be seen as questionable when evaluated against established academic norms of authorship.[58] Moreover, studies have indicated that fully open data sources often lead to reduced credit for data providers in terms of authorship.[3] Additionally, there is a temporal gap between data sharing and publication, and typically, only principal investigators receive group coauthorship recognition, sidelining junior researchers who contribute substantially to the project and operate under stringent publication pressures during their short tenure in a lab. Finally, certain journals now disallow group authorship entirely. In conclusion, reconsidering the mechanisms that drive data sharing and the associated rewards is needed.

Recommendations and resources for motivational aspects

As recently summarised by the Global Research Council, a broader and more inclusive research framework needs new ways of assessing and rewarding research and innovation.[59] More specifically, the need for a restructured credit system to encourage data sharing is becoming increasingly apparent.[3,60] At its core, this system must establish an enduring connection between researchers and the original datasets. Because research data acquisition is typically a collaborative effort involving multiple team members, equitable credit allocation for acquired and shared data is essential. An ideal framework would facilitate personalised and enduring associations between researchers (eg, using their ORCID) and datasets (eg, using a dataset identifier); however, datasets that are extensively used in numerous high-impact publications should carry greater credit weight than obscure collections with little use within a credit system capable of assessing contributions on the basis of data's scientific value. Thus, designing such a system becomes imperative. Consequently, an additional linkage needs to be established between publication identifiers (eg, digital object identifiers) and data identifiers. This interlinking of individual researchers, datasets, and scientific publications facilitates traceability and quantification of dataset use across various publications, enabling credits to be attributed to individual researchers.

These credits could subsequently be made accessible to the broader scientific community, funding bodies, and academic institutions, supplementing the assessment of individual scientists' scientific contributions beyond their own publications. The integration of this new credit system into the criteria for selection and evaluation by academic centres and funding agencies will be pivotal to its effectiveness. Importantly, such a system also provides a mechanism to acknowledge the contributions of junior researchers and scientific support personnel (eg, study nurses, scientific coordinators, and technicians) for their indispensable roles in data collection, in a direct and meaningful manner. Above all, we contend that the implementation of such a credit system would substantially expedite the progress of the open science revolution.

Limitations

Although our survey provided important empirical insights into the perspective of neuroscientists working in the field of neurological and psychiatric disorders with respect to data sharing, our approach has several limitations. The survey request was distributed among ECIB members and principal investigators of associated consortia only, resulting in a small number of responses. This approach might have induced a selection bias and might also limit the generalisability of the survey results; however, the study distribution was intentionally designed in a way to increase the certainty of reaching the intended target group (ie, principal investigators of large European consortia with extensive experience in neuroimaging and presumably a substantial interest in data sharing). To increase the knowledge gained, future surveys should aim to reach a wider audience by using established open science channels (eg, Neurostars or the BIDS mailing list). Furthermore, as some survey questions were optional, the number of repsonses differed per question, and no single question received 95 responses, despite there being 95 survey respondents overall. The result that all survey respondents are willing to share data in the absence of legal or technical constraints should be interpreted with caution. This might reflect a form of conformity bias to meet the implicit expectations of the scientific peer group.

Conclusion

In sum, political action is needed. We call for a harmonised application of the GDPR across EU countries, particularly considering the sharing of brain imaging data and funding specifically dedicated to data sharing. Moreover, data sharing needs to be sufficiently incentivised in the current research system to foster sharing. Therefore, we suggest the development of a system that makes data count by recognising that data sharing is a highly valuable contribution to the scientific and clinical communities.

Contributors

Funding acquisition: KA, FB, SFC, SM, TvE. Conceptualisation: KA, FB, SFC, GC, EDW, ED, GBF, KG, MAI, JJ, SM, H-JM, CP, DP, PR, TvE, LEMW. Literature search: KA, FB, SFC, GBF, KG, SM, H-JM, BS, TvE. Survey design: FB, GC, EDW, ED, GBF, KG, MAI, SM, H-JM, DP, PR, TvE, LEMW. Data curation: KG. Data analysis: KG, TvE. Data interpretation: FB, SFC, SC, GC, ED, GBF, KG, SM, H-JM, WO, DP, PR, TvE, LEMW. Project administration: KA, FB, EDW, KG, SM, H-JM.

Supervision: FB, SFC, KG, SM, H-JM. Visualisation of the figures: FB, KG, SM, H-JM. Writing (original draft): FB, SFC, M-EC, KG, JJ, SM, H-JM, BS, TvE. Writing (review and editing): KA, FB, SFC, SC, GC, M-EC, ED, GBF, KG, MAI, JJ, SM, H-JM, WO, CP, DP, PR, BS, TvE, LEMW.

Declaration of interests

H-JM is supported by the Dutch Heart Foundation (03-004-2020-T049), by the Eurostars-2 joint programme with cofunding from the EU Horizon 2020 Research and Innovation Programme (ASPIRE E!113701), provided by the Netherlands Enterprise Agency, and by the EU Joint Program for Neurodegenerative Disease Research, provided by the Netherlands Organisation for Health Research and Development and Alzheimer Nederland (DEBBIE JPND2020-568-106). FB is supported by the National Institute for Health and Care Research biomedical research centre at University College London Hospitals. He is part of the steering committee or Data Safety Monitoring Board member for Biogen, Merck, Alzheimer's Therapeutic Research Institute and Alzheimer's Clinical Trials Consortium, and Prothena. He is a consultant for Roche, Celltrion, Rewind Therapeutics, Merck, IXICO, Jansen, and Combinostics, has research agreements with Merck, Biogen, GE Healthcare, and Roche, and is a cofounder and shareholder of Queen Square Analytics. SC is supported by the EU Horizon 2020 Research and Innovation Programme under grant agreement number 945539 (HBP SGA3). GC has received research support from the EU Horizon 2020 Research and Innovation Programme under grant agreement number 667696, support from Fondation d'entreprise MMA des Entrepreneurs du Futur, Fondation Alzheimer, Programme Hospitalier de Recherche Clinique, Agence Nationale de la Recherche, Région Normandie, Association France Alzheimer et maladies apparentées, Fondation Vaincre Alzheimer, Fondation Recherche Alzheimer, and Fondation pour la Recherche Médicale (all to Institut National de la Santé et de la Recherche Médicale), and personal fees from Institut National de la Santé et de la Recherche Médicale (salary), Eisai, and Fondation Alzheimer. SM received speaker honoraria from GE Healthcare, Eli Lilly, and Life Molecular Imaging. PR acknowledges support by Virtual Research Environment at Charité Berlin – a node of EBRAINS Health Data Cloud, the EU Horizon Europe programme Horizon EBRAINS2.0 (101147319), Virtual Brain Twin (101137289), EBRAINS-PREP 101079717, AISN – 101057655, EBRAIN-Health101058516, Digital Europe TEF-Health 101100700, EU H2020 Virtual Brain Cloud 826421, Human Brain Project SGA2 785907; Human Brain Project SGA3 945539, ERC Consolidator 683049; German Research Foundation SFB 1436 (project ID 425899996); SFB 1315 (project ID 327654276); SFB 936 (project ID 178316478); SFB-TRR 295 (project ID 424778381); SPP Computational Connectomics RI 2073/6-1, RI 2073/10-2, RI 2073/9-1; DFG Clinical Research Group BECAUSE-Y 504745852, PHRASE Horizon EIC grant 101058240; Berlin Institute of Health & Foundation Charité, Johanna Quandt Excellence Initiative; and ERAPerMed Pattern-Cog2522FSB904. BS is supported by the María de Maeztu Unit of Excellence (Institute of Neurosciences, University of Barcelona) CEX2021-001159-M Ministry of Science, Innovation and Universities and by Generalitat de Catalunya 2021SGR0080 and the CERCA Programme/Generalitat de Catalunya. LEMW is supported by MultiPark, a strategic research area at Lund University, an National Institutes of Health grant (R01-AG070952), an Alzheimerfonden project grant (AF-980872), and a Vetenskapsradet project grant (2022-00900). TvE is a civil servant of North Rhine-Westphalia and employee of the University Hospital of Cologne, Germany. In the last 2 years, he has received honoraria, stipends, or speaker fees from the Lundbeck Foundation, Gain Therapeutics, Orion Pharma, Lundbeck Pharma, Atheneum, and the International Movement Disorders Society. He receives materials from Life Molecular Imaging and Lilly Pharma. He owns stocks of the corporations NVIDIA, Microsoft, and IBM. Multiple unrelated research projects are currently supported by the German Research Foundation. All other authors declare no competing interests.

Acknowledgments

The work of the European Cluster for Imaging Biomarkers, particularly the costs for the online survey conducted via SurveyMonkey and the consensus meeting, was supported by the European Brain Council in the framework of the European Brain Research Area project, which has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement number 825348.

References

1 European Federation of Academies of Sciences and Humanities. The European code of conduct for research integrity. June 2023. https://allea.org/code-of-conduct/ (accessed June 21, 2023).
2 Milham MP, Craddock RC, Son JJ, et al. Assessment of the impact of shared brain imaging data on the scientific literature. *Nat Commun* 2018; **9:** 2818.
3 Nature. Time to recognize authorship of open data. *Nature* 2022; **604:** 8.
4 Carrillo MC, Bain LJ, Frisoni GB, Weiner MW. Worldwide Alzheimer's disease neuroimaging initiative. *Alzheimers Dement* 2012; **8:** 337–42.
5 European Research Executive Agency. Open science. https://rea.ec.europa.eu/open-science_en (accessed March 15, 2024).
6 Paret C, Unverhau N, Feingold F, et al. Survey on open science practices in functional neuroimaging. *Neuroimage* 2022; published online May 17. https://doi.org/10.1016/j.neuroimage.2022.119306.
7 European Commission. She figures 2021: gender in research and innovation: statistics and indicators. Luxembourg: Publications Office of the European Union, 2021.
8 Stewart S, Rinke EM, McGarrigle R, et al. Pre-registration and Registered Reports: a primer from UKRN. *OSF Preprints* 2020; published online Oct 30. https://doi.org/10.31219/osf.io/8v2n7 (preprint).
9 Kerr NL. HARKing: hypothesizing after the results are known. *Pers Soc Psychol Rev* 1998; **2:** 196–217.
10 Simmons JP, Nelson LD, Simonsohn U. False-positive psychology: undisclosed flexibility in data collection and analysis allows presenting anything as significant. *Psychol Sci* 2011; **22:** 1359–66.
11 Rosenthal R. The file drawer problem and tolerance for null results. *Psychol Bull* 1979; **86:** 638–41.
12 Flannery J. fMRI preregistration template. May 6, 2018. https://osf.io/6juft/ (accessed Nov 9, 2023).
13 Center for Open Science. Registered Reports: peer review before results are known to align scientific values and practices. https://www.cos.io/initiatives/registered-reports (accessed Nov 9, 2023).
14 DeHaven A. Preregistration: a plan, not a prison. May 23, 2017. https://www.cos.io/blog/preregistration-plan-not-prison (accessed Nov 9, 2023).
15 Zuo X-N, Xu T, Milham MP. Harnessing reliability for neuroscience research. *Nat Hum Behav* 2019; **3:** 768–71.
16 Nichols TE, Das S, Eickhoff SB, et al. Best practices in data analysis and sharing in neuroimaging using MRI. *Nat Neurosci* 2017; **20:** 299–303.
17 Bell LC, Suzuki Y, van Houdt PJ, Sourbron S, Mutsaerts HJ. The road to the ISMRM OSIPI: a community-led initiative for reproducible perfusion MRI. *Magn Reson Med* 2024; **91:** 1740–42.
18 Shukla-Dave A, Obuchowski NA, Chenevert TL, et al. Quantitative imaging biomarkers alliance (QIBA) recommendations for improved precision of DWI and DCE-MRI derived biomarkers in multicenter oncology trials. *J Magn Reson Imaging* 2019; **49:** e101–21.
19 Smith EE, Biessels GJ, De Guio F, et al. Harmonizing brain magnetic resonance imaging methods for vascular contributions to neurodegeneration. *Alzheimers Dement* 2019; **11:** 191–204.
20 Hendrix JA, Finger B, Weiner MW, et al. The worldwide Alzheimer's disease neuroimaging initiative: an update. *Alzheimers Dement* 2015; **11:** 850–59.
21 Karakuzu A, Appelhoff S, Auer T, et al. qMRI-BIDS: an extension to the brain imaging data structure for quantitative magnetic resonance imaging data. *Sci Data* 2022; **9:** 517.
22 Layton KJ, Kroboth S, Jia F, et al. Pulseq: a rapid and hardware-independent pulse sequence prototyping framework. *Magn Reson Med* 2017; **77:** 1544–52.
23 Karakuzu A, Biswas L, Cohen-Adad J, Stikov N. Vendor-neutral sequences and fully transparent workflows improve inter-vendor reproducibility of quantitative MRI. *Magn Reson Med* 2022; **88:** 1212–28.
24 Hansen MS, Sørensen TS. Gadgetron: an open source framework for medical image reconstruction. *Magn Reson Med* 2013; **69:** 1768–76.
25 Poldrack RA, Markiewicz CJ, Appelhoff S, et al. The past, present, and future of the brain imaging data structure (BIDS). *arXiv* 2023; published online Sept 11. https://doi.org/10.48550/arXiv.2309.05768 (preprint).

26 Pernet CR, Appelhoff S, Gorgolewski KJ, et al. EEG-BIDS, an extension to the brain imaging data structure for electroencephalography. *Sci Data* 2019; **6:** 103.

27 Niso G, Gorgolewski KJ, Bock E, et al. MEG-BIDS, the brain imaging data structure extended to magnetoencephalography. *Sci Data* 2018; **5:** 180110.

28 Norgaard M, Matheson GJ, Hansen HD, et al. PET-BIDS, an extension to the brain imaging data structure for positron emission tomography. *Sci Data* 2022; **9:** 65.

29 Gorgolewski KJ, Alfaro-Almagro F, Auer T, et al. BIDS apps: improving ease of use, accessibility, and reproducibility of neuroimaging data analysis methods. *PLoS Comput Biol* 2017; **13:** e1005209.

30 Maumet C, Nichols BN, Flandin G, et al. Standardized reporting of neuroimaging results with NIDM in SPM, FSL and AFNI. June, 2015. https://www.researchgate.net/publication/281012078_Standardized_reporting_of_neuroimaging_results_with_NIDM_in_SPM_FSL_and_AFNI (accessed Dec 19, 2023).

31 BIDS Contributors. The brain imaging data structure. https://bids-specification.readthedocs.io/en/stable/ (accessed June 21, 2023).

32 Blair R, Contributing Members of BIDS Community. BIDS Validator v1.14.5. https://bids-standard.github.io/bids-validator/ (accessed June 21, 2023).

33 Appelhoff S, Oostenveld R, Blair R, et al. Bids-standard / bids-examples. https://github.com/bids-standard/bids-examples (accessed June 21, 2023).

34 Wilkinson MD, Dumontier M, Aalbersberg IJ, et al. The FAIR Guiding Principles for scientific data management and stewardship. *Sci Data* 2016; **3:** 160018.

35 Wilkinson MD, Dumontier M, Jan Aalbersberg I, et al. Addendum: the FAIR Guiding Principles for scientific data management and stewardship. *Sci Data* 2019; **6:** 6.

36 Towards European Health Data Space. Report on secondary use of health data through European case studies. Feb 28, 2022. https://tehdas.eu/app/uploads/2022/08/tehdas-report-on-secondary-use-of-health-data-through-european-case-studies-.pdf (accessed June 21, 2023).

37 Berlin Institute of Health at Charité. BIH/Charité Virtual Research Environment. https://www.bihealth.org/en/translation/network/digital-medicine/bihcharite-virtual-research-environment (accessed Oct 24, 2023).

38 EBRAINS. Integrate and share your data. https://www.ebrains.eu/data/share-data/share-data-process (accessed June 21, 2023).

39 Ienca M, Fins JJ, Jox RJ, et al. Towards a governance framework for brain data. *Neuroethics* 2022; **15:** 20.

40 Eke DO, Bernard A, Bjaalie JG, et al. International data governance for neuroscience. *Neuron* 2022; **110:** 600–12.

41 Klingner CM, Denker M, Grün S, et al. Research data management and data sharing for reproducible research-results of a community survey of the German National Research Data Infrastructure Initiative Neuroscience. *eNeuro* 2023; **10:** ENEURO.0215-22.2023.

42 White T, Blok E, Calhoun VD. Data sharing and privacy issues in neuroimaging research: opportunities, obstacles, challenges, and monsters under the bed. *Hum Brain Mapp* 2022; **43:** 278–91.

43 EUR-Lex. Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing Directive 95/46/EC (General Data Protection Regulation) (Text with EEA relevance). May 4, 2016. https://eur-lex.europa.eu/eli/reg/2016/679/oj (accessed May 6, 2024).

44 Vlahou A, Hallinan D, Apweiler R, et al. Data sharing under the General Data Protection Regulation: time to harmonize law and research ethics? *Hypertension* 2021; **77:** 1029–35.

45 Peloquin D, DiMaio M, Bierer B, Barnes M. Disruptive and avoidable: GDPR challenges to secondary research uses of data. *Eur J Hum Genet* 2020; **28:** 697–705.

46 Eiss R. Confusion over Europe's data-protection law is stalling scientific progress. *Nature* 2020; **584:** 498–99.

47 The European Parliament and the Council of the EU. Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing Directive 95/46/EC (General Data Protection Regulation). May 4, 2016. https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:32016R0679 (accessed April 10, 2024).

48 Il garante per la protezione dei dati personali. Provvedimento del 7 aprile 2022 [9774819]. April 7, 2022. https://www.garanteprivacy.it/web/guest/home/docweb/-/docweb-display/docweb/9774819 (accessed June 21, 2023).

49 Vokinger KN, Stekhoven DJ, Krauthammer M. Lost in anonymization – a data anonymization reference classification merging legal and technical considerations. *J Law Med Ethics* 2020; **48:** 228–31.

50 Eke D, Aasebø IE, Akintoye S, et al. Pseudonymisation of neuroimages and data protection: increasing access to data while retaining scientific utility. *Neuroimage Rep* 2021; **1:** 100053.

51 Jwa A, Koyejo O, Poldrack R. Demystifying the risk of reidentification in neuroimaging data – a technical and regulatory analysis. *OSF Preprints* 2023; published online Sept 29. https://doi.org/10.31219/osf.io/73mwd (preprint).

52 Schwarz CG, Kremers WK, Wiste HJ, et al. Changing the face of neuroimaging research: comparing a new MRI de-facing technique with popular alternatives. *Neuroimage* 2021; **231:** 117845.

53 Gao P, Dong H-M, Liu S-M, et al. A Chinese multi-modal neuroimaging data release for increasing diversity of human brain mapping. *Sci Data* 2022; **9:** 286.

54 Clunie DA, Flanders A, Taylor A, et al. Report of the Medical Image De-Identification (MIDI) Task Group—best practices and recommendations. *arXiv* 2023; published online Sept 18. https://doi.org/10.48550/arXiv.2303.10473 (preprint).

55 de Sitter A, Visser M, Brouwer I, et al. Impact of removing facial features from MR images of MS patients on automatic lesion and atrophy metrics. *Mult Scler J* 2017; **23:** 226.

56 Bannier E, Barker G, Borghesani V, et al. The Open Brain Consent: informing research participants and obtaining consent to share brain imaging data. *Hum Brain Mapp* 2021; **42:** 1945–51.

57 Lam A. What motivates academic scientists to engage in research commercialization: 'gold', 'ribbon' or 'puzzle'? *Res Policy* 2011; **40:** 1354–68.

58 International Committee of Medical Journal Editors. Defining the role of authors and contributors. https://www.icmje.org/recommendations/browse/roles-and-responsibilities/defining-the-role-of-authors-and-contributors.html (accessed June 21, 2023).

59 Global Research Council. Statement of principles on recognising and rewarding researchers. 2023. https://globalresearchcouncil.org/fileadmin//documents/GRC_Publications/SoP_Recognising_and_Rewarding_Researchers.pdf (accessed Nov 10, 2023).

60 Pierce HH, Dev A, Statham E, Bierer BE. Credit data generators for data reuse. *Nature* 2019; **570:** 30–32.