Current Addiction Reports (2026) 13:38
https://doi.org/10.1007/s40429-026-00746-y

REVIEW

Identifying Impulsivity and Substance use Measures in Neuroimaging Datasets: a Scoping Review and Searchable Dataset Repository for Researchers

Maria I. Perica[1] · Gaylen E. Fronk[1] · Anna E. Kirkland[1] · Deanna Varley[2] · Matthew Sunderland[2] · Louise Mewton[2] · Lindsay M. Squeglia[1]

Received: 17 December 2025 / Accepted: 8 April 2026

Abstract

Purpose of review Impulsivity is a risk factor for problematic substance use. Understanding its neural underpinnings requires large neuroimaging samples, which are difficult and costly to obtain. This scoping review aimed to identify open-access datasets that allow researchers to investigate the neural basis of impulsivity in relation to substance use.

Recent Findings We searched PubMed, Scopus, and major data repositories. 35 openly available datasets with over 200 participants (ranging from 200 to 500,000) were identified that contain (1) impulsivity, (2) substance use, and (3) neuroimaging measures. We compiled this information into a searchable, interactive Shiny-based application to help researchers identify relevant datasets.

Summary Although access to large datasets is improving, more progress is needed to ensure data are findable, accessible, interoperable, and reusable. This review lowers barriers to using open-access data by guiding researchers toward available datasets relevant to the neural basis of impulsivity and substance use, as well as provides recommendations for future data sharing.

Keywords Impulsivity · Substance use · Neuroimaging · Datasets · Open access

Introduction

Impulsivity is a multifaceted construct that can refer to an array of primarily maladaptive behaviors, including difficulty inhibiting responses and hasty decision-making [1, 2]. A substantial body of research has identified high levels of impulsivity as a risk factor for developing problematic substance use and substance use disorders (SUD) [3–7]. This reproducible association between impulsivity and SUDs highlights impulsivity as a promising target for SUD prevention and intervention [8].

✉ Maria I. Perica
perica@musc.edu

1 Department of Psychiatry and Behavioral Sciences, Medical University of South Carolina, Charleston, SC, USA

2 Matilda Centre for Research in Mental Health and Substance Use, University of Sydney, Sydney, Australia

However, the multidimensional nature of impulsivity makes it difficult to study, as distinct neural mechanisms likely underlie its different components; both trait impulsivity and state-dependent fluctuations in inhibitory control contribute to impulsive behavior but these processes may rely on separate neural systems [1]. Since our understanding of the construct-dependent neural mechanisms of impulsivity remains limited, the progress toward developing neuroscience-informed prevention strategies and interventions targeting this core risk factor for SUDs has been slow.

Several barriers exist that impede faster progress in uncovering the neural circuitry leading to impulsivity. First, acquiring neural data using neuroimaging methodologies can be costly and requires specialized equipment and expertise, which is not available to all investigators. Second, even when such resources are available, the timeline from study design to data collection to analysis to publication is lengthy; this timeline only increases when trying to reproduce findings, thus

Published online: 16 April 2026

delaying scientific progress. Further, collecting data from sufficient numbers of participants to be well-powered to detect reproducible effects can be cost-prohibitive. Using large samples is key to identifying reliable brain-behavior associations [9]. Third, although there are some openly available, large datasets that include measures of impulsivity, substance use, and neuroimaging, they can be difficult to locate if they are not deposited in easily findable ways and are therefore underutilized. Finally, even when datasets are located, if metadata is not well-documented, it can take a significant amount of time for researchers to determine if a candidate dataset meets their needs (e.g., includes the specific measures they need, has the age range of interest).

To address these challenges, we conducted a scoping review to identify and catalog available datasets that include (1) impulsivity measures, (2) substance use measures, and (3) neuroimaging measures. We summarize key dataset characteristics of interest to researchers (such as available measures, sample size, age range). Additionally, we discuss the current landscape of available data in relation to the FAIR (Findable, Accessible, Interoperable, and Reusable) principles, which provide a set of descriptive and qualitative guidelines for data sharing, in order to make recommendations for future steps [10]. Further, we provide an online application developed using the Shiny R package that provides a searchable metadata repository to help researchers efficiently identify datasets that include their population or measures of interest [11]. Our goals for this review are to lower barriers for researchers trying to find and access openly available datasets, promote the design of well-powered studies that can clarify the neural mechanisms of impulsivity and their relationship to substance use, and improve future data sharing practices.

Methods

This review was conducted using established guidelines for scoping reviews [12]. We searched for datasets using PubMed, Scopus and individual data repositories including: the National Data Archives, the NIH Brain Development Cohorts (NBDC), NeuroImaging Tools and Resources Collaboratory (NITRC), the Wellcome Trust's report of Landscaping International Longitudinal Datasets, and OpenNeuro. Individual repositories were manually searched in October and November 2025. Searches were split evenly between 3 reviewers (MP, GF, and AK). PubMed and Scopus searches were extracted on October 21, 2025. Studies from PubMed and Scopus were reviewed by one reviewer (MP) in Covidence for inclusion if they were likely to contain relevant datasets [13]. Manuscripts were first screened by abstract and then by full text. Datasets were considered eligible if they:

1. Are available to researchers. For this review, we included datasets with varying access requirements, including payment, to be comprehensive.
2. Are large-scale, which we defined to be datasets with a sample size of at least 200 in the context of the fields of neuroscience and psychology.
3. Included at least one impulsivity measure (e.g., Barratt Impulsiveness Scale), which we defined inclusively to be measures of trait impulsivity as well as cognitive processes contributing to impulsive behavior, such as inhibitory control, as per [7].
4. Included at least one substance use measure, which we defined to be datasets that have any measure of quantity or frequency of substance use or measures assessing problematic substance use (e.g., Alcohol Use Disorders Identification Test).
5. Included at least one type of neuroimaging measure (e.g., structural or functional MRI, EEG).
6. Included human participant data.
7. Were deposited in a data repository or published in a peer-reviewed journal.

For the PubMed and Scopus search, search terms were used to capture datasets that included impulsivity, substance use, and neuroimaging, yielding 122 results (see Supplementary Materials for search terms). From the PubMed and Scopus literature review, a total of 47 studies met criteria for describing a dataset of interest, and from those 47 studies, 11 unique datasets were identified (Fig. 1). These datasets included the Adolescent Brain Cognitive Development Study [14], the HEALthy Brain and Child Development Study [15], the Nathan Kline Institute/Rockland Sample [16], IMAGEN [17], the Two-Decades Brainclinics Archive for Insights into Neurophysiology [18], the Duke Neurogenetics Study [19], the Human Connectome Project Young Adult 1200 [20], the UK Biobank [21], the Enhancing Neuroimaging Genetics through Meta-Analysis Consortium Addiction Working Group [22, 23], the UCLA Consortium for Neuropsychiatric Phenomics [24], and the Consortium on Vulnerability to Externalizing Disorders and Addictions [25]. For manual searching of data repositories, if filtering functionality was available, the first pass to identify datasets filtered on basic characteristics of dataset (e.g., sample size of 200 requirement, basic broad keywords such as "substance use"). Datasets that passed that first pass were then assessed for eligibility more thoroughly to determine whether they met all our criteria. Manual

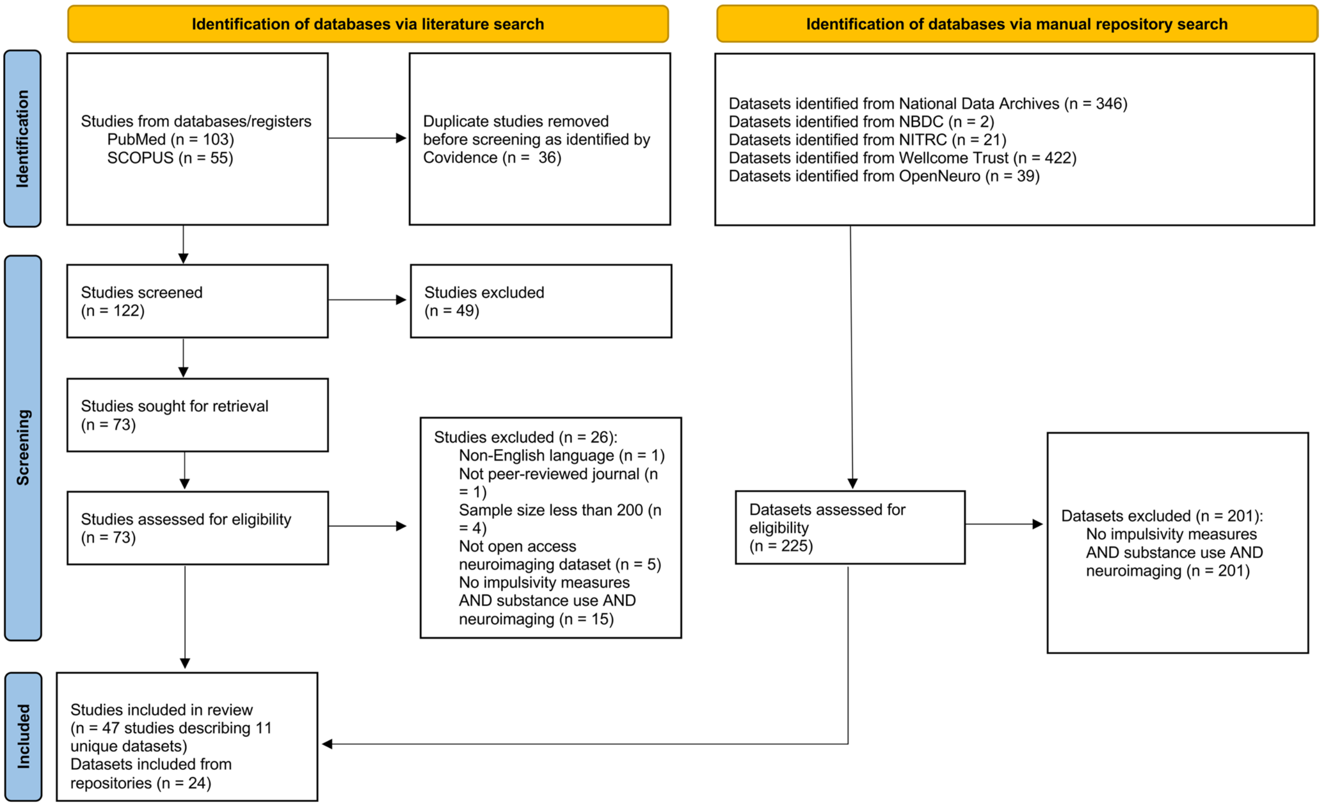

Fig. 1 Adapted PRISMA 2020 flowchart describing the screening and selection of studies from literature and datasets from repository

searching of data repositories yielded an additional 24 datasets [26–46]. If datasets were duplicated between literature search and manual repository search, they were only included in results and metadata resource once.

Using the identified datasets, features of datasets were coded into a spreadsheet with a pre-defined data extraction template by the person reviewing the manuscript or repository, and pulled into R to summarize data, create figures, and build an online metadata resource [R version 4.5.2, R Core Team]. These features included age range of sample, country where data were collected, study design (cross-sectional or longitudinal), impulsivity measures, substance use measures, neuroimaging methods, dataset access link and citation, and a qualitative assessment of how well data adhered to the FAIR principles as described and how well data were documented by available metadata resources (e.g., is data file format explicitly mentioned). Well-known and frequently used measures of impulsivity, substance use, and neuroimaging as identified by the literature [7] and by the research team were coded with the measure's name; if measures existed that were less commonly used or known, they were coded as 'Other' to maximize the datasets we included in our repository. The 'shiny' package was used to create an online metadata resource to help search through available datasets [11].

Results

We identified 35 datasets in total (Fig. 1). Of those datasets, 17 had between 200 and 500 participants, 5 datasets had 500-1,000, 8 had 1,000–5,000, and 5 had over 5,000 (Fig. 2A; range: 200–500,000; total participants across all datasets: 556,299). Regarding study design of the datasets, 14 had a cross-sectional design, 14 had a longitudinal design, and 7 datasets contained both cross-sectional and longitudinal data (Fig. 2B). Most datasets contained data collected in the United States ($n=29$), but there were also 2 datasets that contained data from England and Germany, and 1 dataset that contained data from France, India, Ireland, Scotland, the Netherlands, and Wales (Fig. 2D; note that some datasets contribute to multiple categories). One dataset (ENIGMA) was a consortium of many datasets with data shared from sites across the world.

In terms of the measures contained across datasets, there was little consensus in any category (Fig. 2E and G). The most commonly used impulsivity measures were the Behavioral Inhibition System/Behavioral Activation System (BIS/BAS), the Flanker Task, and delay discounting tasks (Fig. 2E). The most commonly used substance use measures were toxicology measures, the Structured Clinical Interview for DSM Disorders (SCID), and the Alcohol Use Disorders

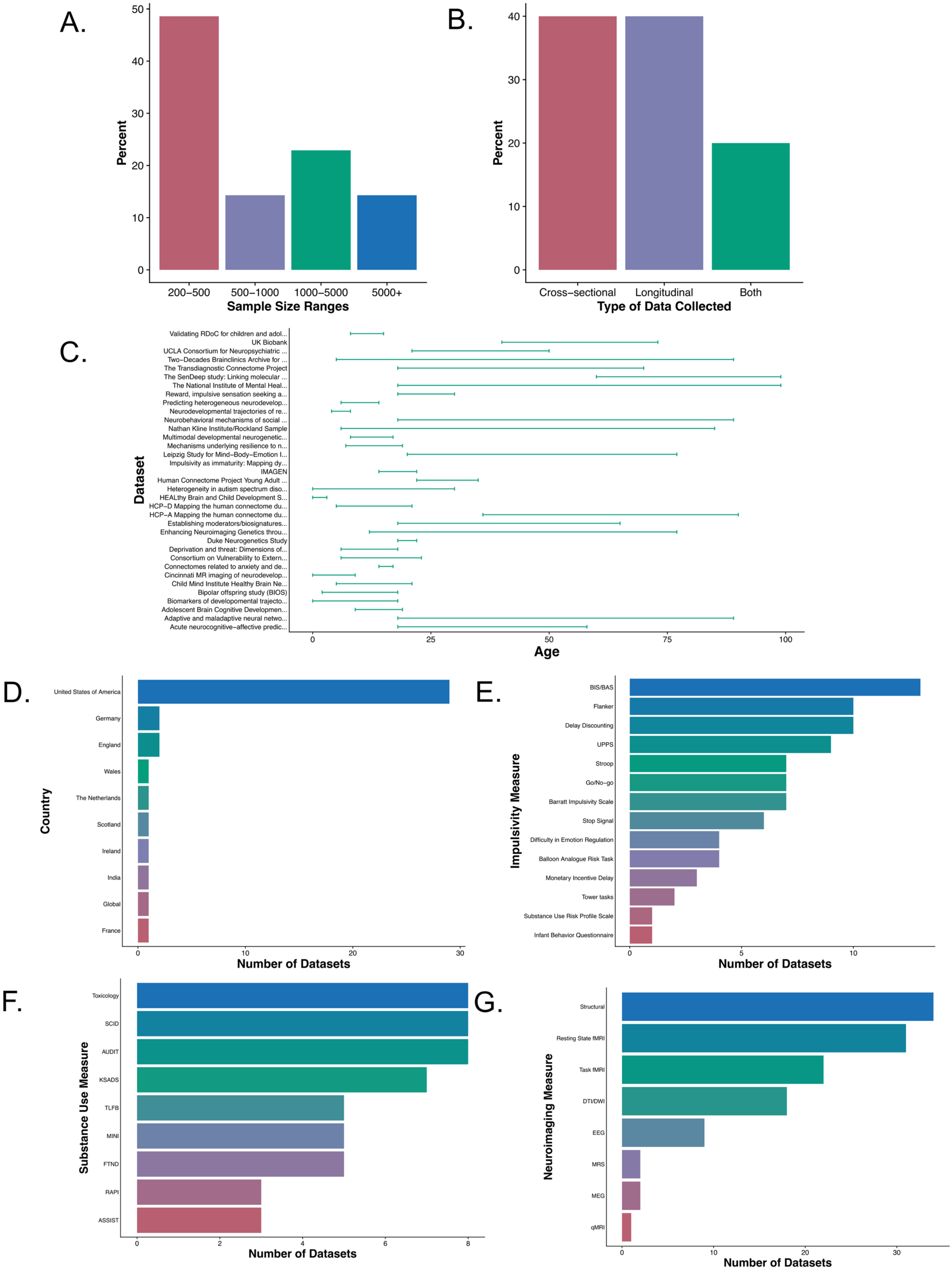
A.
Percent
0
10
20
30
40
50
200–500
500–1000
1000–5000
5000+
Sample Size Ranges
B.
Percent
0
10
20
30
40
Cross–sectional
Longitudinal
Both
Type of Data Collected
C.
Dataset
Validating RDoC for children and adol...
UK Biobank
UCLA Consortium for Neuropsychiatric ...
Two–Decades Brainclinics Archive for ...
The Transdiagnostic Connectome Project
The SenDeep study: Linking molecular ...
The National Institute of Mental Heal...
Reward, impulsive sensation seeking a...
Predicting heterogeneous neurodevelop...
Neurodevelopmental trajectories of re...
Neurobehavioral mechanisms of social ...
Nathan Kline Institute/Rockland Sample
Multimodal developmental neurogenetic...
Mechanisms underlying resilience to n...
Leipzig Study for Mind–Body–Emotion I...
Impulsivity as immaturity: Mapping dy...
IMAGEN
Human Connectome Project Young Adult ...
Heterogeneity in autism spectrum diso...
HEALthy Brain and Child Development S...
HCP–D Mapping the human connectome du...
HCP–A Mapping the human connectome du...
Establishing moderators/biosignatures...
Enhancing Neuroimaging Genetics throu...
Duke Neurogenetics Study
Deprivation and threat: Dimensions of...
Consortium on Vulnerability to Extern...
Connectomes related to anxiety and de...
Cincinnati MR imaging of neurodevelop...
Child Mind Institute Healthy Brain Ne...
Bipolar offspring study (BIOS)
Biomarkers of developomental trajecto...
Adolescent Brain Cognitive Developmen...
Adaptive and maladaptive neural netwo...
Acute neurocognitive–affective predic...
0
25
50
75
100
Age
D.
Country
United States of America
Germany
England
Wales
The Netherlands
Scotland
Ireland
India
Global
France
0
10
20
30
Number of Datasets
E.
Impulsivity Measure
BIS/BAS
Flanker
Delay Discounting
UPPS
Stroop
Go/No–go
Barratt Impulsivity Scale
Stop Signal
Difficulty in Emotion Regulation
Balloon Analogue Risk Task
Monetary Incentive Delay
Tower tasks
Substance Use Risk Profile Scale
Infant Behavior Questionnaire
0
5
10
Number of Datasets
F.
Substance Use Measure
Toxicology
SCID
AUDIT
KSADS
TLFB
MINI
FTND
RAPI
ASSIST
0
2
4
6
8
Number of Datasets
G.
Neuroimaging Measure
Structural
Resting State fMRI
Task fMRI
DTI/DWI
EEG
MRS
MEG
qMRI
0
10
20
30
Number of Datasets

Fig. 2 Description of datasets. **A** Count of datasets that contain different ranges of sample sizes. **B** Count of datasets that contain different types of data. **C** Range of ages contained in each dataset. **D** Count of datasets per country. **E** Count of datasets containing most commonly used impulsivity measures. **F** Count of datasets containing most commonly used substance use measures. **G** Count of datasets containing most commonly used neuroimaging measures. Note: some counts do not sum to the total number of datasets because information was not available for all datasets, and some datasets include multiple data types and therefore appear in more than one category

Identification Test (AUDIT) (Fig. 2F). The most commonly used neuroimaging methods were structural MRI, resting state and task fMRI, followed by diffusion tensor or diffusion-weighted imaging (DTI/DWI). Of our identified datasets, almost all included either structural or functional MRI (34 out of 35 datasets). Fewer included other neural measures such as DTI/DWI ($n=18$), EEG (electroencephalography; $n=9$), magnetoencephalography (MEG; $n=2$), magnetic resonance spectroscopy (MRS; $n=2$), and quantitative MRI ($n=1$).

The majority of datasets (34) were publicly indexed with persistent identifiers (e.g., digital object identifiers), which allow for more permanent and reliable referencing. Most (31) were free to use but required some level of standard authorization to access, which frequently included a project proposal that required approval, in addition to affiliation with an academic institution. Only 1 dataset required payment to access the data, and only 4 datasets were entirely open to download without any barriers. Datasets that were deposited in data repositories (e.g., National Data Archive) required specific data formats and variable description to be deposited and thus were necessarily available in a standard format, such as a CSV file, and information about specific measures included was readily available. However, 3 independently hosted datasets did not describe what format data would be available in, and 4 datasets did not describe what specific measures were available.

Shiny application

To assist researchers in finding and using this data, we built a metadata resource adapted from Varley et al. (2025) [47]. This web-based, interactive Shiny application allows users to view the characteristics of all datasets identified in this review, and filter datasets by their criteria of interest to better identify what datasets may be available for use: https://gaylen-fronk.shinyapps.io/shiny_sud_impulsivity/.

To use this application, a researcher can adjust multiple filtering options on the menu on the lefthand side of the page to filter on dataset features including but not limited to: year of data collection, age of study sample participants, study type (cross-sectional or longitudinal), impulsivity/substance use/neuroimaging measures included, and whether the dataset is accessible (e.g., fully open with no restriction/protocol, open via standard protocol, etc.). Datasets that match the filtered upon criteria then appear on the righthand side of the page with a link for researchers to easily find them.

Discussion

In this review, we aimed to identify and record datasets that are available to researchers interested in understanding the neural mechanisms of impulsivity, and to provide a useful resource that can lower the barriers to finding this data. We identified 35 publicly available datasets, which included a total of 556,299 participants. The datasets vary in their levels of openness and accessibility, with 31 accessible via application, 3 available for download without application, and 1 available via application and payment. We summarized the key characteristics of these datasets and provide a searchable metadata resource that offers a convenient tool that allows users to easily search for and identify datasets that match criteria of interest.

It was notable that we were able to identify 35 openly accessible datasets that include data from 200 or more participants, with 21 containing at least some longitudinal data. This is a positive step toward ensuring that all researchers can conduct reproducible and rigorous science, as compiling and summarizing these datasets in a single resource reduces barriers to data access and encourages data reuse while improving interoperability across studies. However, it was often difficult to find information about which datasets contained repeated measures or multiple timepoints and which did not. Going forward, it is necessary to more clearly describe this key feature of datasets so other researchers easily know whether the shared data can answer their research questions of interest.

Further, while sharing data from large samples provide an invaluable resource for shifting the field further from non-representative convenience samples and closer to population neuroscience [48], there are some limitations to the data being shared. For example, the majority of datasets identified (approximately 86%) contain data collected in the United States, thus limiting global generalizability. This highlights a need for increased data collection and sharing from a broader array of countries to enable replication across diverse samples, and support data harmonization efforts that integrate data from multiple regions, ultimately creating large datasets that are more representative of a global population.

Interestingly, there was little consistency in what measures were used to operationalize each of our constructs of interest (impulsivity, substance use, neural measures).

This inconsistency may have to do with the complexity of thoroughly capturing these constructs themselves, but the heterogeneity across measures creates barriers to data harmonization efforts, requiring researchers to have to take additional steps to combine across datasets [49, 50]. By developing a consensus set of measures for these constructs, as exemplified by the Phenotypes and eXposures (PhenX) Toolkit [51], data harmonization could be simplified.

Regarding neuroimaging data, it has become more common over the past few decades to openly share neuroimaging data. Repositories such as OpenNeuro assist scientists in sharing neuroimaging data in a way that best aligns with the FAIR principles [52]. Compared to other types of data, neuroimaging data are particularly costly to collect and often limited to universities and medical centers that have the necessary neuroimaging equipment. Therefore, by sharing these data more openly, the field democratizes neuroimaging data and makes them more widely accessible to scientists that would not otherwise be able to work with these data, enabling more science and new insights. As aforementioned, the most common neuroimaging data that were shared in our identified datasets were structural and functional MRI, while other types of data, such as EEG or MRS were less commonly available. This highlights a need for increased sharing of other types of neural data as well as data not found at all in this review such as positron emission tomography (PET) data. The sharing of a more diverse array of neuroimaging data would enable a more comprehensive understanding of neural systems.

Although the existence of these datasets is significant progress in the realm of data sharing within the field, adherence to best practices for data sharing as described by the FAIR principles remains variable[10]. Regarding findability, most identified datasets were publicly indexed with persistent identifiers (typically digital object identifiers or DOIs). However, it should be noted that this review may be limited in its ability to identify datasets that are less findable and do not have persistent identifiers in our search of PubMed, Scopus, and data repositories. Therefore, it is possible this review did not identify every shared dataset possible that includes our variables of interest. This limitation underscores the importance of emphasizing findability when sharing data.

Another barrier to being able to use shared data is the accessibility of the data. Only 3 identified datasets are fully open-access, with few-to-no barriers. This may partially be due to the sensitive nature of data collected; some datasets required authorization in part to protect any sensitive or identifying participant information and ensure researchers agreed to responsible conduct of research. This degree of accessibility is still in line with FAIR principles for data sharing, and many datasets had clear, standardized communication procedures detailed on their websites. However, some datasets required more extensive approval of the proposed project itself by the data collecting team, as well as manuscript authorship credit. While it is understandable that researchers would have concerns about receiving credit for the substantial amount of labor invested in collecting the data, this provides added degrees of human intervention that can significantly impede a researcher's access to the data and is not recommended by the FAIR principles [10, 53]. Finally, many identified datasets would benefit from improving the findability and open accessibility of their data dictionaries and other metadata so researchers may more easily ascertain exactly what measures are included in the dataset to determine whether the dataset is of interest. As an example, the Adolescent Brain Cognitive Development and HEALthy Brain Cognitive Development studies require all researchers to obtain authorization to access the primary data files, but the names of the measures themselves are accessible through a user-friendly online portal that anyone can explore. Further, data shared through the National Data Archive are required to be accompanied by publicly accessible information about the specific measures and items in the dataset.

Additionally, information required to assess interoperability was often either difficult to locate or unavailable. For example, datasets were not consistently accompanied by information noting whether they were available in standard, widely-used formats such as CSV files, and it was often unclear whether variables were labelled using consistent, easily understandable and discipline-appropriate coding and naming conventions. Likewise, it was often unknown whether any naming conventions were maintained across related datasets (such as multiple waves of a longitudinal study) to facilitate straightforward linking.

Information relevant to reusability was frequently incomplete as well. Very few datasets clearly reported basic provenance details such as when data collection started and finished. Many dataset descriptions relied on information from initial grant proposals rather than the actual characteristics of the data that were ultimately collected. This lack of accurate, up-to-date metadata is particularly problematic in contexts where timing matters or where data collection differs from planned protocols in meaningful ways. For example, the ABCD study provides a model of documenting the timing of data collection and procedure changes when they had to shift to virtual collection quickly when the COVID-19 pandemic hit, and in-person neuroimaging visits were postponed [54]. When datasets do not clearly document timing of data collection, researchers are unable consider potential effects of event or cohort.

Overall, many shared datasets would benefit from substantially clearer and more comprehensive metadata, including accurate descriptions, detailed provenance information (e.g., when, where, and how the data were collected), data processing and cleaning information, and usage licenses to improve other researchers' ability to easily reuse the shared data. Further, while data repositories provide useful platforms for researchers to store and share data, they often lack features to filter by key variables or data characteristics, making it more difficult to locate data of interest. We therefore hope that the Shiny app we developed will be useful both as a tool that provides this functionality and as a potential template for demonstrating what features of datasets are of relevance and interest to researchers, and what information is required to describe datasets in a way that makes them easy to understand and reuse.

Conclusion

Our review synthesizes available, open-access datasets that include measures of impulsivity, substance use, and neuroimaging. We catalogue these datasets, describe their key features, and provide a searchable metadata resource to help reduce barriers researchers may encounter when trying to find and access relevant data. We also highlight areas where future data-sharing practices can be improved. By increasing data sharing and improving data sharing practices, we can accelerate the speed at which the field is able to uncover novel insights. Data sharing can also reduce research costs for many scientists, ensure the investment of federal funding goes further by producing more insights from the same data, maximize the contributions of research participants, and enhance rigor and reproducibility [53]. By consolidating information on accessible datasets, this review contributes to ongoing efforts to facilitate research on the neural mechanisms underlying impulsivity in substance use, thereby accelerating progress in this important area of research.

Key References

- Verdejo-Garcia A, Albein-Urios N. Impulsivity traits and neurocognitive mechanisms conferring vulnerability to substance use disorders. Neuropharmacology 2021;183:108402. https://doi.org/10.1016/j.neuropharm.2020.108402.
    - This review outlines the multifaceted complexity of the construct of 'impulsivity.'
- Wilkinson MD, Dumontier M, Aalbersberg IjJ, Appleton G, Axton M, Baak A, et al. The FAIR Guiding Principles for scientific data management and stewardship. Sci Data 2016;3:160018. https://doi.org/10.1038/sdata.2016.18.
    - The seminal paper describing the FAIR guiding principles for data sharing.
- Markiewicz CJ, Gorgolewski KJ, Feingold F, Blair R, Halchenko YO, Miller E, et al. The OpenNeuro resource for sharing of neuroscience data. eLife 2021;10:e71774. https://doi.org/10.7554/eLife.71774.
    - This paper describes the OpenNeuro data repository as an example of a data sharing platform that adheres to the FAIR principles, and discusses aspects specific to neuroimaging (e.g., BIDS).
- Poldrack RA, Gorgolewski KJ. Making big data open: data sharing in neuroimaging. Nat Neurosci 2014;17:1510–7. https://doi.org/10.1038/nn.3818.
    - This review outlines the current state of data sharing in neuroimaging and addresses potential researcher concerns and future challenges.

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1007/s40429-026-00746-y.

Acknowledgements We would like to thank Emily Brennan, MLIS of MUSC Libraries for her assistance with the PubMed and Scopus literature search.

Author contributions M.P. wrote the main manuscript and prepared all figures. G.F., D.V., and M.P. wrote code for and built Shiny App. M.P., G.F., and A.K. reviewed literature and data repositories. All authors provided input on structure and content of review, feedback on Shiny app, and provided feedback on and reviewed the manuscript.

Funding Open access funding provided by the Carolinas Consortium. MP is supported by NIDA and NIAAA (T32DA007288, L30AA032673) grants. LS is supported by NIAAA (K24AA031052). AK is supported by NIAAA (K01AA031745).

Data Availability No datasets were generated or analysed during the current study.

Declarations

Competing interests The authors declare no competing interests.

Human and Animal Rights and Informed Consent No animal or human subjects were used in this study.

References

1. de Wit H. Impulsivity as a determinant and consequence of drug use: a review of underlying processes. Addict Biol. 2009;14:22–31. https://doi.org/10.1111/j.1369-1600.2008.00129.x.
2. Diagnostic and Statistical Manual of Mental Disorders. American Psychiatric Association Publishing; 2013. https://doi.org/10.1176/appi.books.9780890425596.
3. Dawe S, Loxton NJ. The role of impulsivity in the development of substance use and eating disorders. Neurosci Biobehav Rev. 2004;28:343–51. https://doi.org/10.1016/j.neubiorev.2004.03.007.
4. Green R, Meredith LR, Mewton L, Squeglia LM. Adolescent neurodevelopment within the context of impulsivity and substance use. Curr Addict Rep. 2023;10:166–77. https://doi.org/10.1007/s40429-023-00485-4.
5. Perry JL, Carroll ME. The role of impulsive behavior in drug abuse. Psychopharmacology. 2008;200:1–26. https://doi.org/10.1007/s00213-008-1173-0.
6. Verdejo-García A, Lawrence AJ, Clark L. Impulsivity as a vulnerability marker for substance-use disorders: review of findings from high-risk research, problem gamblers and genetic association studies. Neurosci Biobehav Rev. 2008;32:777–810. https://doi.org/10.1016/j.neubiorev.2007.11.003.
7. Verdejo-Garcia A, Albein-Urios N. Impulsivity traits and neurocognitive mechanisms conferring vulnerability to substance use disorders. Neuropharmacology. 2021;183:108402. https://doi.org/10.1016/j.neuropharm.2020.108402.
8. Zhong G, Chen T, Zhong N, Rezapour T, Haghparast A, Jiang H, et al. Transdiagnostic neuromodulation of impulsivity: current status and future trajectories. Transl Psychiatry. 2025;15:209. https://doi.org/10.1038/s41398-025-03415-2.
9. Marek S, Tervo-Clemmens B, Calabro FJ, Montez DF, Kay BP, Hatoum AS, et al. Reproducible brain-wide association studies require thousands of individuals. Nature. 2022;603:654–60. https://doi.org/10.1038/s41586-022-04492-9.
10. Wilkinson MD, Dumontier M, Aalbersberg I, Appleton G, Axton M, Baak A, et al. The FAIR guiding principles for scientific data management and stewardship. Sci Data. 2016;3:160018. https://doi.org/10.1038/sdata.2016.18.
11. Chang W, Cheng J, Allaire JJ, Sievert C, Schloerke B, Aden-Buie G, et al. shiny: Web Application Framework for R. 2025.
12. Arksey H, O'Malley L. Scoping studies: towards a methodological framework. Int J Soc Res Methodol. 2005;8:19–32. https://doi.org/10.1080/1364557032000119616.
13. Veritas Health Innovation. Covidence systematic review software 2025.
14. Volkow ND, Koob GF, Croyle RT, Bianchi DW, Gordon JA, Koroshetz WJ, et al. The conception of the ABCD study: From substance use to a broad NIH collaboration. Dev Cogn Neurosci 2017;32:4–7. https://doi.org/10.1016/j.dcn.2017.10.002.
15. Volkow ND, Gordon JA, Bianchi DW, Chiang MF, Clayton JA, Klein WM, et al. The HEALthy brain and child development study (HBCD): NIH collaboration to understand the impacts of prenatal and early life experiences on brain development. Dev Cogn Neurosci. 2024;69:101423. https://doi.org/10.1016/j.dcn.2024.101423.
16. Nooner KB, Colcombe SJ, Tobe RH, Mennes M, Benedict MM, Moreno AL, et al. The NKI-Rockland sample: a model for accelerating the pace of discovery science in psychiatry. Front Neurosci. 2012;6:152. https://doi.org/10.3389/fnins.2012.00152.
17. Mascarell Maričić L, Walter H, Rosenthal A, Ripke S, Quinlan EB, Banaschewski T, et al. The IMAGEN study: a decade of imaging genetics in adolescents. Mol Psychiatry. 2020;25:2648–71. https://doi.org/10.1038/s41380-020-0822-5.
18. van Dijk H, van Wingen G, Denys D, Olbrich S, van Ruth R, Arns M. The two decades brainclinics research archive for insights in neurophysiology (TDBRAIN) database. Sci Data 2022;9:333. https://doi.org/10.1038/s41597-022-01409-z.
19. Baranger DAA, Few LR, Sheinbein DH, Agrawal A, Oltmanns TF, Knodt AR, et al. Borderline personality traits are not correlated with brain structure in two large samples. Biol Psychiatry Cogn Neurosci Neuroimaging. 2020;5:669–77. https://doi.org/10.1016/j.bpsc.2020.02.006.
20. Elam JS, Glasser MF, Harms MP, Sotiropoulos SN, Andersson JLR, Burgess GC, et al. The human connectome project: a retrospective. Neuroimage. 2021;244:118543. https://doi.org/10.1016/j.neuroimage.2021.118543.
21. Sudlow C, Gallacher J, Allen N, Beral V, Burton P, Danesh J, et al. UK biobank: an open access resource for identifying the causes of a wide range of complex diseases of middle and old age. PLoS Med. 2015;12:e1001779. https://doi.org/10.1371/journal.pmed.1001779.
22. Mackey S, Kan K-J, Chaarani B, Alia-Klein N, Batalla A, Brooks S, et al. Genetic imaging consortium for addiction medicine: from neuroimaging to genes. Prog Brain Res. 2016;224:203–23. https://doi.org/10.1016/bs.pbr.2015.07.026.
23. Thompson PM, Jahanshad N, Schmaal L, Turner JA, Winkler AM, Thomopoulos SI, et al. The Enhancing NeuroImaging Genetics through Meta-Analysis Consortium: 10 years of global collaborations in human brain mapping. Hum Brain Mapp. 2021;43:15–22. https://doi.org/10.1002/hbm.25672.
24. Gorgolewski KJ, Durnez J, Poldrack RA. Preprocessed consortium for neuropsychiatric phenomics dataset. F1000Res. 2017;6:1262. https://doi.org/10.12688/f1000research.11964.2.
25. Sharma E, Jacob P, Murthy P, Jain S, Varghese M, Jayarajan D, et al. Consortium on Vulnerability to Externalizing Disorders and Addictions (cVEDA): a developmental cohort study protocol. BMC Psychiatry. 2020;20:2. https://doi.org/10.1186/s12888-019-2373-3.
26. Alexander LM, Escalera J, Ai L, Andreotti C, Febre K, Mangone A, et al. An open resource for transdiagnostic research in pediatric mental health and learning disorders. Sci Data. 2017;4:170181. https://doi.org/10.1038/sdata.2017.181.
27. Bagdasarov A, Roberts K, Bréchet L, Brunet D, Michel CM, Gaffrey MS. Spatiotemporal dynamics of EEG microstates in four- to eight-year-old children: age- and sex-related effects. Dev Cogn Neurosci. 2022;57:101134. https://doi.org/10.1016/j.dcn.2022.101134.
28. Bounoua N, Spielberg JM, Sadeh N. Dimensions of childhood maltreatment and neural circuits of affective inhibitory control: An fMRI study. Brain Res Bull. 2025;230:111507. https://doi.org/10.1016/j.brainresbull.2025.111507.
29. Burt SA, Klump KL, Vazquez AY, Shewark EA, Hyde LW. Identifying Patterns of Youth Resilience to Neighborhood Disadvantage. Res Hum Dev. 2021;18:181–96. https://doi.org/10.1080/15427609.2021.1935607.
30. Chopra S, Cocuzza CV, Lawhead C, Ricard JA, Labache L, Patrick LM, et al. The Transdiagnostic Connectome Project: an open dataset for studying brain-behavior relationships in psychiatry. Sci Data. 2025;12:923. https://doi.org/10.1038/s41597-025-04895-z.

31. Corley RP, Reynolds CA, Wadsworth SJ, Rhea S-A, Hewitt JK. The Colorado Twin Registry: 2019 Update. Twin Res Hum Genet. 2019;22:707–15. https://doi.org/10.1017/thg.2019.50.
32. Golding, Pembrey, Jones, Team TAS. ALSPAC–The Avon Longitudinal Study of Parents and Children. Paediatric and Perinatal Epidemiology 2001;15:74–87. https://doi.org/10.1046/j.1365-3016.2001.00325.x.
33. Gozdas E, Holland SK, Altaye M, CMIND Authorship Consortium. Developmental changes in functional brain networks from birth through adolescence. Hum Brain Mapp 2019;40:1434–44. https://doi.org/10.1002/hbm.24457.
34. Hubbard NA, Siless V, Frosch IR, Goncalves M, Lo N, Wang J, et al. Brain function and clinical characterization in the Boston adolescent neuroimaging of depression and anxiety study. Neuroimage Clin. 2020;27:102240. https://doi.org/10.1016/j.nicl.2020.102240.
35. Juttukonda MR, Li B, Almaktoum R, Stephens KA, Yochim KM, Yacoub E, et al. Characterizing cerebral hemodynamics across the adult lifespan with arterial spin labeling MRI data from the Human Connectome Project-Aging. Neuroimage. 2021;230:117807. https://doi.org/10.1016/j.neuroimage.2021.117807.
36. Liuzzi MT, Harb F, Petranu K, Huggins AA, Webb EK, Fitzgerald JM, et al. The Dichotomy of Threat and Deprivation as Subtypes of Childhood Maltreatment: Differential Functional Connectivity Patterns of Threat and Reward Circuits in an Adult Trauma Sample. Biol Psychiatry Cogn Neurosci Neuroimaging. 2024;9:227–34. https://doi.org/10.1016/j.bpsc.2023.10.001.
37. Lorenzo EC, Figueroa JE, Demirci DA, El-Tayyeb F, Huggins BJ, Illindala M, et al. Unraveling the association between major depressive disorder and senescent biomarkers in immune cells of older adults: a single-cell phenotypic analysis. Front Aging. 2024;5:1376086. https://doi.org/10.3389/fragi.2024.1376086.
38. Nugent AC, Thomas AG, Mahoney M, Gibbons A, Smith JT, Charles AJ, et al. The NIMH intramural healthy volunteer dataset: a comprehensive MEG, MRI, and behavioral resource. Sci Data. 2022;9:518. https://doi.org/10.1038/s41597-022-01623-9.
39. Pan LA, Goldstein TR, Rooks BT, Hickey M, Fan JY, Merranko J, et al. The relationship between stressful life events and axis I diagnoses among adolescent offspring of probands with bipolar and non-bipolar psychiatric disorders and healthy controls: the Pittsburgh bipolar offspring study (BIOS). J Clin Psychiatry. 2017;78:e234-43. https://doi.org/10.4088/JCP.15m09815.
40. Ressa HJ, Newman BT, Jacokes Z, McPartland JC, Kleinhans NM, Druzgal TJ, et al. Widespread associations between behavioral metrics and brain microstructure in ASD suggest age mediates subtypes of ASD. Imaging Neurosci (Camb). 2025;3:IMAG.a.144. https://doi.org/10.1162/IMAG.a.144.
41. Sanders AFP, Harms MP, Kandala S, Marek S, Somerville LH, Bookheimer SY, et al. Age-related differences in resting-state functional connectivity from childhood to adolescence. Cereb Cortex. 2023;33:6928–42. https://doi.org/10.1093/cercor/bhad011.
42. Schmidt NL, Brooker RJ, Carroll IC, Gagne JR, Luo Z, Planalp EM, et al. Longitudinal research at the interface of affective neuroscience, developmental psychopathology, health and behavioral genetics: findings from the Wisconsin twin project. Twin Res Hum Genet. 2019;22:233–9. https://doi.org/10.1017/thg.2019.55.
43. Schumer MC, Bertocci MA, Aslam HA, Graur S, Bebko G, Stiffler RS, et al. Patterns of neural network functional connectivity associated with mania/hypomania and depression risk in 3 independent young adult samples. JAMA Psychiatry. 2024;81:167–77. https://doi.org/10.1001/jamapsychiatry.2023.4150.
44. Trivedi MH, McGrath PJ, Fava M, Parsey RV, Kurian BT, Phillips ML, et al. Establishing moderators and biosignatures of antidepressant response in clinical care (EMBARC): rationale and design. J Psychiatr Res. 2016;78:11–23. https://doi.org/10.1016/j.jpsychires.2016.03.001.
45. Vannucci A, Fields A, Bloom PA, Camacho NL, Choy T, Durazi A, et al. Probing the content of affective semantic memory following caregiving-related early adversity. Dev Sci. 2024;27:e13518. https://doi.org/10.1111/desc.13518.
46. Way, Browning, Christopher R., Wagner, Dylan D., Ford, Jodi L., Boettner, Bethany, Bai, Ping. Structural and functional MRI dataset from the Adolescent Health and Development in Context (AHDC) study in Columbus, Ohio. 2025. https://doi.org/10.18112/openneuro.ds005896.v1.0.0.
47. Varley D, Henry A, Halladay J, Baillie A, Keyes K, Slade T, et al. Identifying psychological distress data available in nationally representative surveys: a scoping review and case study of Australian surveys. Soc Psychiatry Psychiatr Epidemiol. 2025. https://doi.org/10.1007/s00127-025-02981-6.
48. Falk EB, Hyde LW, Mitchell C, Faul J, Gonzalez R, Heitzeg MM, et al. What is a representative brain? Neuroscience meets population science. Proc Natl Acad Sci U S A. 2013;110:17615–22. https://doi.org/10.1073/pnas.1310134110.
49. Gibbons RD, Perraillon MC, Kim JB. Item response theory approaches to harmonization and research synthesis. Health Serv Outcomes Res Method. 2014;14:213–31. https://doi.org/10.1007/s10742-014-0125-x.
50. Sunderland M, Olsen N, Visontay R, Chapman C, Mewton L, Stapinski L, et al. "One metric to rule them all": a common metric for symptoms of depression and generalized anxiety in adolescent samples. Clin Psychol Sci. 2024;12:147–60. https://doi.org/10.1177/21677026231168564.
51. Hamilton CM, Strader LC, Pratt JG, Maiese D, Hendershot T, Kwok RK, et al. The PhenX toolkit: get the most from your measures. Am J Epidemiol. 2011;174:253–60. https://doi.org/10.1093/aje/kwr193.
52. Markiewicz CJ, Gorgolewski KJ, Feingold F, Blair R, Halchenko YO, Miller E, et al. The OpenNeuro resource for sharing of neuroscience data. Elife. 2021;10:e71774. https://doi.org/10.7554/eLife.71774.
53. Poldrack RA, Gorgolewski KJ. Making big data open: data sharing in neuroimaging. Nat Neurosci. 2014;17:1510–7. https://doi.org/10.1038/nn.3818.
54. Feldstein Ewing SW, Dash GF, Thompson WK, Reuter C, Diaz VG, Anokhin A, et al. Measuring retention within the adolescent brain cognitive development (ABCD)SM study. Dev Cogn Neurosci. 2022;54:101081. https://doi.org/10.1016/j.dcn.2022.101081.