ANNUAL REVIEWS

Annual Review of Developmental Psychology

Poverty, Brain Development, and Mental Health: Progress, Challenges, and Paths Forward

Christopher S. Monk[1,2] and Felicia A. Hardi[1]

[1]Department of Psychology, University of Michigan, Ann Arbor, Michigan, USA; email: csmonk@umich.edu, felhardi@umich.edu

[2]Survey Research Center, Michigan Neuroscience Institute, and Department of Psychiatry, University of Michigan, Ann Arbor, Michigan, USA

Annu. Rev. Dev. Psychol. 2023. 5:309–30

First published as a Review in Advance on July 12, 2023

The *Annual Review of Developmental Psychology* is online at devpsych.annualreviews.org

https://doi.org/10.1146/annurev-devpsych-011922-012402

Keywords

adversity, neuroimaging, developmental psychopathology, replicability, reproducibility

Abstract

Poverty is associated with changes in brain development and elevates the risk for psychopathology in childhood, adolescence, and adulthood. Although the field is rapidly expanding, there are methodological challenges that raise questions about the validity of current findings. These challenges include the interrelated issues of reliability, effect size, interindividual heterogeneity, and replicability. To address these issues, we propose a multipronged approach that spans short-, medium-, and long-term solutions, including changes to data pipelines along with more comprehensive data acquisition of environment, brain, and mental health. Additional suggestions are to use open science approaches, more robust statistical analyses, and replication testing. Furthermore, we propose increased integration between advanced analytical approaches using large samples and neuroscience models in intervention research to enhance the interpretability of findings. Collectively, these approaches will expand the application of neuroimaging findings and provide a foundation for eventual policy changes designed to improve conditions for children in poverty.

Contents

INTRODUCTION

Growing economic inequality (Saez & Zucman 2014) and deteriorating mental health among youth (Keyes et al. 2019) have increased the urgency to better understand how individual differences in brain development explain the links between environmental conditions and risk for mental disorders. Poverty has a profound influence on children's day-to-day experience through a myriad of processes including increased parental stressors that hamper caregiving practices, reduced community social support, and elevated neighborhood crime (Brooks-Gunn & Duncan 1997, Conger et al. 2002, Evans 2004, McLoyd 1998). Moreover, poverty and its associated sequelae diminish human potential and increase the risk for chronic conditions, including mental disorders, that can last a lifetime (Costello et al. 2003, Lazzarino et al. 2013). Although recent findings indicate that poverty is robustly linked to brain development (Tomasi & Volkow 2021), new insights concerning the reliability of brain measures and the small effect sizes linking brain to phenotypes suggest that we are far from understanding how poverty alters brain development and increases the risk for psychopathology. This is further complicated by the fact that human brain development is highly heterogeneous, and there are wide individual differences in the experience of poverty and its effects on mental health outcomes.

The goal of the present review is to assess the state of the field, characterize methodological barriers, and identify paths forward. The review is divided into three sections. In the first section, we briefly review the existing literature on poverty, brain development, and mental health. The second section focuses on methodological challenges in conducting this research. Finally, the third section offers possible paths forward for increasing the reliability, replicability, and interpretability of the research. These paths include changes to phenotyping, neuroimaging, and analyses. Some challenges, such as improvements to the data processing pipeline, are relatively easy to address.

In contrast, other challenges, including the heterogeneity of brain, clinical symptoms, and links between brain and symptoms, pose more fundamental problems for the field. As part of this section, we provide an overview of large data sets that may be used to answer relevant questions and discuss the strengths and limitations of these data sets. Finally, we discuss possible approaches for the field to increase interpretability and relevance for policy and clinical applications.

OVERVIEW OF THE EXISTING LITERATURE

Evidence from longitudinal population-based and birth-cohort studies has demonstrated that poverty is a potent stressor for child and adolescent mental health via many associated risk factors. To establish causal effects, intervention studies have demonstrated that alleviating poverty can improve child development and well-being. For instance, quasi-experiments and studies on earned income tax credits have both shown that providing financial assistance to families can help improve children's academic achievement and reduce child behavioral problems (Akee et al. 2010, Dahl & Lochner 2012). Similarly, experimental studies that provide opportunities to move to more affluent neighborhoods and increase social support (Axinn et al. 2022, Chetty et al. 2016) or studies that provide conditional and unconditional cash transfers to families (Duncan et al. 2011, 2014) have found that increased access to resources can reduce risks for subsequent mental health problems in children.

At the same time, many children who grew up in poverty do not develop mental health problems, suggesting that there could be individual differences in poverty effects that could have neurobiological bases (Garmezy 1993, Masten et al. 2021). In particular, the effects of poverty on individual developmental trajectories could differ through variations in biological sensitivity to the environment or the interaction between environment and biology (Boyce & Ellis 2005, Jensen et al. 2017). These findings led to efforts to understand the neural mechanisms that underlie the link between poverty and mental health, which could help inform targets for prevention and intervention efforts.

Much progress has been made on this front. Socioeconomic status has been associated with differences in whole-brain volume and cortical surface area, white matter structural connectivity and fiber density, and brain function both during rest and while performing in-scanner tasks (Brito & Noble 2014, Hackman & Farah 2009, Hair et al. 2015, Rakesh & Whittle 2021). Recently, in one impressive set of findings, Tomasi & Volkow (2021) systematically examined cross-sectional associations between measures of socioeconomic status (SES) and structural magnetic resonance imaging (MRI) measures in the Adolescent Brain Cognitive Development (ABCD) study and found large, reproducible effect sizes in the association between family income and cortical thickness. Moreover, they randomly split the full sample into discovery and validation subsamples ($N = 3{,}892$ each) so that reproducibility could be established. Relative to cortical thickness, the effect sizes were smaller for cortical volume, suggesting that cortical thickness may be a more promising metric when examining SES–brain associations. These findings suggest that a standard measure of poverty, specifically family income, may be reproducibly associated with brain measures in a cross-sectional design.

While most studies have primarily focused on cross-sectional associations, increasing efforts have been made to examine longitudinal trajectories of brain development as a function of poverty exposure (Evans 2004, Hanson et al. 2013, Tooley et al. 2021). Furthermore, identifying specificity across both different experiences of poverty and different neuroimaging modalities (Goetschius et al. 2020b, Hardi et al. 2022, McLaughlin & Sheridan 2016) has increased understanding of the mechanisms and processes through which poverty relates to brain and socioemotional development. Finally, important work is underway using a randomized controlled trial to examine the effects of poverty reduction on brain development and cognition (Troller-Renfree et al. 2022).

METHODOLOGICAL CHALLENGES

Across the research field that examines poverty, brain development, and mental health, there are multiple challenges that can undermine findings. In this section, issues regarding specific methods as well as challenges in linking measures across multiple levels are discussed.

Measuring Poverty

Despite established links, obstacles in measuring poverty have hindered progress and yielded contradictory results. In particular, it has been argued that common proxies of poverty, such as income and education, are fundamentally flawed as they fail to capture important factors such as government subsidies, wealth accumulation, geographical differences, and educational opportunities (Duncan & Magnuson 2003, Mueller & Parcel 1981). Furthermore, there are vast differences in the experience of poverty, even when family income is comparable. For example, children with similar economic circumstances may experience different levels of parenting and school quality; such factors may modulate the effects of poverty on child development (Shavers 2007). Poverty's impact on development is also likely to vary across different developmental stages as children's economic needs change and they gain a more nuanced understanding of social class. Moreover, SES often changes over time, further complicating our ability to capture poverty's effects on development.

Neuroimaging

Accumulating evidence indicates that variability in processing and analytical methods across studies decreases the likelihood of reproducible results. Variations in neuroimaging data acquisition and processing methods (e.g., preprocessing pipelines, different scanning parameters during data acquisition, brain segmentations or motion correction methods) can introduce biases in the resulting brain estimates (Bloom et al. 2022, Carp 2013, Li et al. 2022, Poldrack et al. 2017). Moreover, Li and colleagues (2022) found that agreement among different pipelines is especially poor when used to process data with low reliability, which is a known limitation of many neuroimaging data with a short scan duration. Neuroimaging pipeline issues are also compounded by problems of statistical power. Specifically, neuroimaging studies are often underpowered due to small sample sizes (Button et al. 2013), which in turn leads to inflated effect sizes and replication challenges (Marek et al. 2022, Schulz et al. 2022).

Linking Brain and Phenotype

Recent papers have raised serious concerns about the field's ability to link brain structure and function to mental health symptoms. Specifically, analyses using large data sets, including the ABCD study, UK Biobank, and the Human Connectome Project, examined the associations between brain function and structures with mental health symptoms, cognitive assessments, and personality variables (Chen et al. 2022, Marek et al. 2022, Ooi et al. 2022, Winter et al. 2022). Functional MRI (fMRI) had stronger associations with behavior than structural MRI, and cognitive measures had stronger associations with brain measures than mental health and personality measures (Ooi et al. 2022). Moreover, while associations were identified, the effect sizes, particularly with mental health measures, were small. This is especially alarming given that the sample sizes of many neuroimaging studies are modest, and thus the studies are severely underpowered for detecting these small effect sizes. As a result, many findings with smaller sample sizes may not replicate.

What might account for the small effect sizes? Four possibilities relate to reliability and heterogeneity, and it is likely that most or all of these possibilities contribute to the small effects.

First, the test-retest reliability for diagnostic and symptom measures for relevant mental health conditions and symptoms is often modest (Regier et al. 2013, Tiego et al. 2023). As detailed in an exceptional analysis by Nikolaidis et al. (2022), if one phenotypic measure has poor reliability, the joint reliability of the mental health measure and the biological measure will be compromised, and this will severely attenuate the effect size and degrade replicability between the biological measure and phenotype. Moreover, an item response theory analysis of the Child Behavior Checklist (CBCL), the primary symptom measure of mental health in the ABCD study, showed that the CBCL had poor reliability at the low levels of psychopathology where most participants resided (Tiego et al. 2023). Thus, in combination with the issues uncovered in statistical modeling (Nikolaidis et al. 2022), poor reliability of phenotypic measurement could account for the small effect sizes in some recent large-scale studies as well.

Second, whereas structural MRI and sensorimotor tasks with fMRI typically have excellent test-retest reliability (Elliott et al. 2020), other fMRI protocols, such as rest during shorter acquisition periods and emotion-based tasks with standard fMRI procedures, have modest-to-low test-retest reliability (Elliott et al. 2020, Gratton et al. 2020). Thus, just as poor phenotype reliability impacts brain–behavior study replicability, suboptimal test-retest reliability with brain imaging will severely limit brain–behavior study replicability as well.

Third, mental health conditions are heterogeneous. Specifically, not only can nonoverlapping symptoms lead to a given diagnosis (Gratton et al. 2020), but also the causal mechanisms, including the relative proportion of genetic and environmental contributions along with the developmental timing, may vary within a set of similar symptoms and conditions (Feczko et al. 2019). Given these dimensions of phenotypic heterogeneity, it is clear that there is not a straightforward relationship between brain and phenotype.

Fourth, seminal work from a project called the Midnight Scan Club, in which over 20 hours of high-quality data from 10 individuals were acquired, has revealed tremendous interindividual heterogeneity in brain function (Gordon et al. 2017, Gratton et al. 2018). These findings suggest that the relationship between brain and behavior may be even more complex than previously believed. Moreover, in light of recent concerns regarding small effect sizes, it was proposed that there is likely tremendous heterogeneity in the relationship between brain and behavior across individuals (Bandettini et al. 2022). This means that one specific brain pattern may not necessarily lead to a given outcome. Instead, akin to the concept of multifinality (Cicchetti & Rogosch 1996), there may be various patterns across individuals that contribute to the same symptom category or mental disorder. If confirmed, this would increase the challenge of identifying reliable correlates and point to a pressing need to move away from univariate analyses and toward approaches that consider heterogeneity.

Beyond their possible causes, the issue of small effect sizes themselves in linking brain and behavior poses problems for clinical translation. In particular, small effects make it unlikely that current brain-based approaches can be used to develop biomarkers for improved diagnosis and treatments for mental disorders (Gratton et al. 2020, Schmaal 2022). This is because the interindividual heterogeneity and poor reliability that contribute to small effects make neuroimaging imprecise at the individual level (Gratton et al. 2020).

Conclusion

In summary, although promising findings show that measures of poverty and brain indices yield reliable results, there are specific concerns around neuroimaging measures as well as approaches that link neuroimaging and mental health. These issues are limiting the potential for neuroimaging to lead to translational breakthroughs in diagnosis and treatment. The following section identifies directions to improve replicability.

PATHS FORWARD TO INCREASE REPLICABILITY

Introduction

Despite recent findings showing that the association between poverty, specifically family income, and brain development has large and reproducible effect sizes (Tomasi & Volkow 2021), substantial challenges and opportunities for improvement exist broadly in the field. Problems such as small effect sizes, heterogeneity, and reliability are not unique to this area of research. For example, although the effect sizes of lead exposure and attention problems are small (Goodlad et al. 2013), they are still important and relevant for policy. Moreover, as pointed out in Marek et al. (2022), the small brain–behavior associations based on the ABCD Study, the Human Connectome Project, and the UK Biobank are larger than genome-wide association studies linking genes to behavior. Nevertheless, the major goals of neuroimaging are to contribute to diagnosis as well as to identify best courses of treatment for patients. These goals require individual-level precision, which is presently not possible. Below we discuss approaches to increase reproducibility and reliability in research involving the measurement of poverty, brain development, and mental health. Some of these solutions will take years to implement, while others are relatively straightforward. **Figure 1** outlines a summary of changes that can be implemented immediately, in the medium term, and over the long term. The terms noted in **Figure 1** are further discussed in the sections below.

Open Science

As persuasively and broadly argued in recent years, incorporating open science approaches, including preregistration, data sharing, and code sharing, curtails problems related to scientific practices, such as *p*-hacking, analytic flexibility, and underpowered studies, along with human error (Lindsay 2015). In this field of research, especially with the increasing prevalence of big data where thousands of variables can be examined and tested, continuing the use of open science will increase the rigor of the work and increase the likelihood of replication.

Economic and Poverty Measures

To improve the reproducibility of poverty measures, it is important to include indices other than self-report questionnaires. Adding sources such as administrative data across generations or census tract data (Chetty et al. 2016, Milne et al. 2022) offers a more holistic measure of poverty. Furthermore, consideration should be given to co-occurring measures of inequality such as immigration or refugee status as well as interpersonal and structural racism, which present unique challenges to individuals that could limit access to resources and worsen the severity of poverty (Viruell-Fuentes et al. 2012). In addition, the inclusion of experiential data, such as material hardship, would better capture the lived experience of poverty (Gershoff et al. 2007). To integrate these multilevel data, a structural equation approach can be employed to capture the shared variance between different poverty measures, or combined scores can be computed to capture cumulative effects of poverty exposure (Mueller & Parcel 1981). Another approach is to test whether specific measures of poverty have unique contributions for brain development, such as material hardship (Hardi et al. 2022), household instability (Hardi et al. 2023b), and neighborhood-level poverty (Taylor et al. 2020) over and above family income or other metrics of poverty. Evidence for both the potential accumulated effect and specificity of poverty experience is important for informing targeted interventions.

Finally, more work is needed to test the developmental timing effects of poverty. The effects of poverty may vary depending on the developmental stage of children, as their perception of social

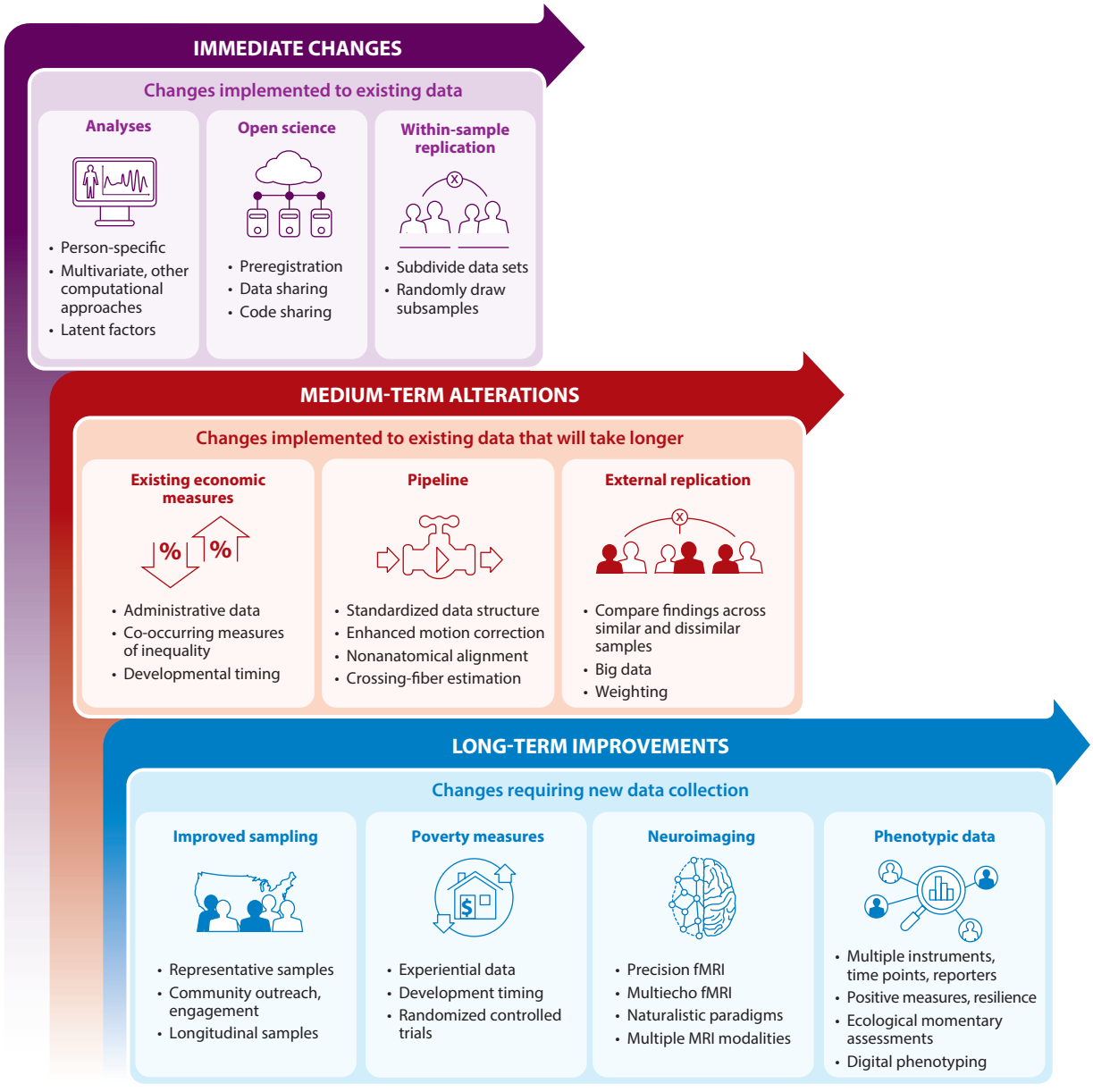

Figure 1

Toward a reproducible neuroscience of poverty, resilience, and mental health. Changes are broken down by the time horizon in which they can be implemented, showing immediate, medium-term, and long-term changes for improving research. Abbreviations: fMRI, functional MRI; MRI, magnetic resonance imaging.

class changes considerably with age (McLoyd 2019). Moreover, parents may be better equipped to shield younger children from certain aspects of poverty, resulting in differing experiences of poverty as children grow older. Additionally, changes in income and resources across time could also have important implications. Experimental research found that moving to a more affluent

neighborhood could have a greater positive impact for younger children (Chetty et al. 2016). At the same time, changes during later childhood are linked to worse outcomes (Natl. Inst. Child Health Hum. Dev. Early Child Care Res. Netw. 2005). This suggests that changes occurring during late childhood may be more egregious even though evidence indicates a sensitive period for brain development occurs during early childhood when the brain is rapidly developing (Giedd et al. 1999, Hensch & Bilimoria 2012).

Neuroimaging

Over the years, procedures have been developed to improve the replicability of neuroimaging data. These procedures, which include improvements to data collection, data pipeline processing, and analysis, are described below. Although not a panacea, implementation of these procedures is expected to lead to substantial improvements in the replicability of neuroimaging data.

Data collection. To mitigate the problems of low reliability and heterogeneity in neuroimaging, acquiring larger quantities of data per subject is important. To that end, precision fMRI and dense sampling, an approach whereby hours of data are collected per participant, have yielded marked increases in reliability (Birn et al. 2013; Gratton et al. 2018, 2020; Noble et al. 2017). However, this approach is not feasible for younger participants who are less likely to be willing or able to undergo hours of data collection. In addition, due to cost constraints, precision fMRI makes collecting data on a large sample impractical. A more promising approach is multiecho fMRI. In standard fMRI acquisition protocols, an excitation pulse is applied to the brain and a single picture (or single echo) of the whole brain is acquired. In multiecho fMRI, multiple pictures of the whole brain are acquired following each pulse (Kundu et al. 2013, Poser et al. 2006). Whereas the physiological signal of interest decays, noise from motion and other sources does not. Therefore, multiecho is more effective than single echo at isolating the relevant fMRI signal, resulting in more reliable and reproducible findings compared with single-echo acquisition (Cohen et al. 2021). Moreover, 10 min of data collection with multiecho yielded superior test-retest reliability compared with 30 min of single echo, and, thus, this procedure may help point to a way forward for conducting precision MRI without a long acquisition time that is not feasible for children (Lynch et al. 2020). Of note, although findings from these procedures were first published in 2006 (Poser et al. 2006), multiecho was initially challenging to implement in many MRI centers, but technical advances have now made it broadly available (Elliott et al. 2020).

Beyond the amount and form of fMRI acquisition, the procedures that the participant undergoes (i.e., the brain state) (Greene et al. 2018) are also important for linking brain to phenotype. Since it was first described in 1995 (Biswal et al. 1995), resting-state functional connectivity, spontaneous fMRI signal in the absence of a task (Fox et al. 2005), has come to play a prominent role in fMRI research (Finn 2021). Advantages to resting-state functional connectivity are that instructions are minimal and participants will not make errors that will force them to be removed from the analysis. More importantly, unlike in task-based fMRI where there are many variations on similar tasks, most resting-state data procedures are the same; thus data from different sites may be easily combined (Biswal et al. 2010). This is a major advantage when pooling large data sets from different groups. Nevertheless, recent findings show that task-based functional connectivity explained the variance of the target phenotype more than threefold compared with the resting state (Greene et al. 2018). Relatedly, naturalistic paradigms, such as movie watching, yielded functional connectivity estimates that predicted cognitive and emotional function better than the resting state (Finn & Bandettini 2021). Why might task-based functional connectivity outperform the resting state? The unconstrained nature of the resting state allows for more unaccounted-for variability in the data, possibly from changes in attention and arousal (Greene

et al. 2018). In contrast, tasks constrain the mental state and possibly tap into the brain processes that are relevant to the phenotype of interest (Greene et al. 2018).

Another potential approach for increasing reliability is combining multiple MRI modalities (e.g., task-based fMRI, resting state fMRI, diffusion MRI, anatomical MRI), which may yield more comprehensive measurement of brain development and provide more replicable findings linking brain to behavior. To date, results utilizing this approach have been somewhat promising, but they require further testing. Whereas analyses that combined different modalities predicted cognitive function better (Jiang et al. 2020, Schulz et al. 2022), other studies reported that the inclusion of diffusion and anatomical data did not provide more predictive power on top of functional connectivity measures (Ooi et al. 2022). Nevertheless, additional predictive power has been shown to be achieved through multimodal fusion, which allows for more complete use of data from each modality (Sui et al. 2020) and has been successfully used to predict cognition in schizophrenia (Sui et al. 2018). Finally, identifying common connectivity patterns within individuals across rest and tasks has also shown promise for increasing reliability (Elliott et al. 2019), but further work is necessary to examine using this approach with mental health measures.

Data pipeline improvements. Possible biases introduced by variations in data acquisition and preprocessing pipelines can be mitigated by using standardized methods and open science approaches that could improve replication efforts. Standardized data structures, such as the Brain Imaging Data Structure (Gorgolewski et al. 2016), provide a generalized way to organize and describe neuroimaging data, improving data interpretation and sharing within or across research groups. In addition, enhanced motion correction provides more reliable neuroimaging markers (for comparisons across commonly used denoising techniques for fMRI data, see, e.g., Parkes et al. 2018); however, stringent criteria of motion can also lead to participant exclusions, which can lead to biases in the results (Gard et al. 2023). This is especially important given evidence showing that children from more disadvantaged backgrounds move more in the scanner (Cosgrove et al. 2022). Thus, stringent motion correction could inadvertently limit the representation of children of interest in research on poverty. As such, while more research is needed to establish best practices in quality control for motion, all studies should establish detailed quality control procedures, report any differences in findings associated with motion, and include motion parameters as a part of the analytic or study design (Parkes et al. 2018).

New methods have been developed to address other limitations in neuroimaging analyses. For instance, a known problem in neuroimaging research is that most commonly used methods assume anatomical and functional coherence across individuals, even though brain function is highly individualized (Gordon et al. 2017, Gratton et al. 2018). Thus, mapping individual brain images onto a normalized standard space may obfuscate important individual differences in brain function. To mitigate this problem, nonanatomical alignment methods, such as hyperalignment (Haxby et al. 2011), have been developed to map functional data into a shared high-dimensional space such that neural regions are aligned functionally as opposed to anatomically. This would overcome the significant source of variability stemming from anatomical differences across participants and has been shown to increase effect sizes of associations with measures of cognitive function (Feilong et al. 2021). While this method has mostly been tested using functional data collected during naturalistic viewing, there is promise for its application for fMRI data in general.

For structural connectivity modeling using diffusion data, new advances in techniques have been developed as an alternative to conventional tractography methods, which are prone to problems in estimating white matter in regions of crossing fibers (Tournier et al. 2012). Techniques that utilize methods optimized for crossing-fiber estimation, such as probabilistic tractography or fixel-based analysis, are effective in establishing a more reliable marker for white matter

development (Dhollander et al. 2021). Additional methods (e.g., Bayesian multitensor model selection) are currently being tested to optimize such estimation even further (Seider et al. 2022), thus advancing our ability to track white matter fibers that facilitate communication across the brain.

Analytic approaches. Advances in analytics also provide exciting ways forward that would improve linkages among poverty, brain, and phenotypes. Heterogeneity in brain function indicates that common statistical approaches that average brain function across participants may not be optimal in capturing individual differences (Elliott et al. 2021, Fair et al. 2021, Feczko et al. 2019). Similarly, as discussed above, psychopathology may manifest in individuals in person-specific ways. These factors suggest that reliable understanding of how poverty affects brain development and mental health may require person-specific analyses (Feczko et al. 2019, Gratton et al. 2020, Molenaar 2004) that model individualized brain networks. Personalized brain network approaches, such as fingerprinting (Fair et al. 2021, Finn et al. 2015, Miranda-Dominguez et al. 2014, Mueller et al. 2013), that establish an individual neural signature have been used to examine person-specific differences in functional networks. Another person-centered approach is to identify specific neuroimaging-based subtypes by clustering individuals within a sample using their neural patterns. This method, which simultaneously minimizes similarities between individuals within the same group and maximizes differences between subgroups, can be used to classify individuals for prediction of risk factors or prospective outcomes. For example, network models using functional connectivity identified adolescents who were susceptible to anxiety six years later (Hardi et al. 2023a) and those who were exposed to violence early in life (Goetschius et al. 2020a). A similar approach has also been taken to distinguish clinically depressed from nondepressed individuals on the basis of brain function (Drysdale et al. 2017, Price et al. 2017). Future investigations may wish to combine these person-specific approaches with an examination of developmental trajectories. Such an approach is comparable to charting developmental trajectories of individuals on pediatric growth charts (Marquand et al. 2019, Rutherford et al. 2022). Similar to other analyses, this method aims to reduce reliance on group averages when making comparisons across patients or healthy groups.

Improving reproducibility for prediction of individual differences also requires methods that better index the reliability of brain data as well as improve existing analytic approaches. For instance, stable variability in the functional signal across a large number of trials may be necessary for meaningful phenotypic prediction of individual differences and can be better estimated using trial-level variability (Chen et al. 2021, Månsson et al. 2022). Furthermore, mounting evidence indicates that methods that are based on multivariate analyses increase power in predictive analyses. As demonstrated by multiple studies (e.g., Kragel et al. 2021, Spisak et al. 2022), multivariate network analyses that aggregate information across the connectome can enhance statistical power and improve the ability to detect effects with smaller samples. For instance, approaches such as canonical correlation analysis (Hardoon et al. 2004, Wang et al. 2020) that maximize the linear relationship between a set of predictors and outcomes can achieve reproducible results with a sample size of 300–1,000 (Tiego et al. 2023). This approach has been shown to be successful in identifying specific network patterns that map onto dimensions of psychopathology (Xia et al. 2018).

To achieve a similar goal of identifying reproducible and reliable neural markers that could account for heterogeneity, computational models that use both unsupervised and supervised machine learning methods are also becoming increasingly popular (Bzdok & Meyer-Lindenberg 2018, Rutherford 2020). These methods typically include a cross-validation process where models are trained and established using a set of data, which are then used for prediction on an independent test sample. Support vector machine classifiers, for instance, have been used to distinguish between healthy and patient samples using brain features and patterns (for a review, see Orrù

et al. 2012), potentially providing a way to predict disease without behavioral data. Similarly, deep learning neural network algorithms can also be used to develop predictive models for more reliable representation of disorders (Durstewitz et al. 2019).

Phenotyping

The reliability of symptom and diagnostic measures of mental health conditions is modest and often poor. For example, as discussed above, the CBCL may have poor reliability in nonclinical samples (Tiego et al. 2023). As for pediatric diagnostic instruments, the Kiddie Schedule for Affective Disorders and Schizophrenia (K-SADS) was found to have excellent or good interrater and test-retest reliability across mental health conditions, but the sample size was small and the number of patients who participated in the reliability testing was even smaller (Kaufman et al. 1997). Thus, more work is needed to assess the reliability of the K-SADS. However, since the K-SADS typically relies on multiple reporters (i.e., the child and a parent), it may provide reasonable reliability. For adults, the Structured Clinical Interviews for the DSM (SCID), a standard instrument for adult diagnoses, had only modest reliability for categorizing many mental disorders (Shankman et al. 2018). This poor reliability substantially contributes to small effect sizes when linking brain to symptoms, and thus, increasing phenotypic reliability offers marked increases in effect size independent of the reliability of the MRI data (Nikolaidis et al. 2022).

How might the reliability of the mental health measures be increased? First, acquiring data using multiple instruments, multiple time points, and multiple reporters (e.g., self-report, parent report, clinician report) effectively increases reliability (Nikolaidis et al. 2022). Second, generating latent factors from multiple instruments also increases reliability (Tiego et al. 2023). Third, combining a clinician interview approach along with inclusion of symptom severity may increase reliability. For example, Shankman and colleagues (2018) adapted the SCID to assess symptom severity and found that it substantially increased the psychometric properties of the instrument. A similar approach with an instrument such as the K-SADS may also be effective. Fourth, the inclusion of ecological momentary assessment data, which provides self-reported information about mood and experiences throughout the day, may provide rich, contextual information that would further enhance the stability of subjective experiences. Finally, data from participant smartphones, including app use, exercise, sleep, and social interactions, provide a digital phenotype that may complement existing clinical measures (Montag & Quintana 2023). In summary, single questionnaires with modest amounts of data do not provide sufficient reliability. Moving forward, it will be crucial to collect data from multiple time points and multiple informants that will yield more reliable indices of emotional experience.

Big Data Approaches

Over the last 12 years, there has been a proliferation of large, publicly available data sets. The Philadelphia Neurodevelopmental Cohort helped usher in this era by collecting and sharing neuroimaging data on an impressive number of children, adolescents, and young adults ($N = 1{,}145$) (Satterthwaite et al. 2014). Moreover, relevant to the present review, this study was particularly groundbreaking because many of the participants were from low-income backgrounds, a segment that had often been ignored in research. More recently, sample sizes have dramatically expanded. Examples of pediatric studies include the ABCD and Healthy Brain Network (HBN) studies. Along with consortia that curate previously collected data, such as Enhancing Neuro Imaging Genetics Through Meta-Analysis (ENIGMA) and the Autism Brain Imaging Data Exchange (Di Martino et al. 2014, Novak et al. 2012), large-scale studies have transformed the way that brain research is done. No longer do individual labs collect and retain data sets; now, many aspects of the large data sets are shared so that novel research questions can be addressed without additional

data collection. These large-scale studies provide a huge resource for the broader neuroimaging community. Although none of these studies were explicitly created to examine possible effects of poverty on brain development, these large-scale studies can be used to address these questions. Moreover, due to their immense size, these endeavors will play a crucial role in resolving issues of reproducibility. Below we discuss the advantages and disadvantages of specific samples.

Adolescent Brain Cognitive Development. The ABCD study collected data from almost 12,000 9- and 10-year-olds from 21 sites across the United States and is now following the sample by collecting lab-based measurements every year and MRI data every other year (Casey et al. 2018, Garavan et al. 2018). Additionally, the project was thoughtfully sampled, with the demographics roughly approximating US demographics (Garavan et al. 2018). Even though it is not the central focus, the ABCD study is suitable for examining poverty, brain development, and mental health. It is estimated that the ABCD study includes approximately 1,034 children who live below the poverty line with acceptable MRI data and complete income data (Ellwood-Lowe et al. 2021). Furthermore, several indices of poverty were collected at multiple time points, including household income; the Parent Reported Financial Adversity Questionnaire, which assesses whether families have enough money to pay for essentials (Taylor et al. 2020, Tomasi & Volkow 2021); and neighborhood-level poverty measures such as the Area Deprivation Index created using residential addresses (Taylor et al. 2020).

Despite well-sampled poverty measures and good representation of individuals below the poverty line, there are several limitations of the sample for linking brain to phenotype. First, the administration of only two types of mental health assessments, the K-SADS and CBCL, may not be adequate in capturing mental health outcomes. The ABCD study used a self-administered, computerized version of the K-SADS in which the child reported on themselves and the parent reported on the child, and it was not administered by clinically trained individuals. The concordance among the self-administered versions and the clinician version is generally modest across conditions based on the Cohen's kappa, but stronger based on Gwet's AC1 (Townsend et al. 2020), raising questions about the diagnoses. Furthermore, at least following the first wave of data collection, symptom levels based on the CBCL scores were low (Chen et al. 2022, Ooi et al. 2022). This could have contributed to the small effect size in brain–behavior association in the sample (Marek et al. 2022) as the CBCL has poor reliability when used to measure mental health outcome in samples with low levels of psychopathology (Tiego et al. 2023). The CBCL was designed to examine clinical samples and so it may lack the sensitivity to measure individual differences in healthy individuals. Lastly, while longitudinal data spanning across multiple developmental stages will become available in the future, data collection began when participants were 9–10 years of age and so any information acquired about prior developmental periods was based on retrospective accounts, which are less reliable (Hardt & Rutter 2004). Despite these limitations, the ABCD study includes an exceptionally large sample of participants from diverse economic backgrounds who are being followed longitudinally. Thus, this project provides important data that has the potential to yield highly reproducible findings.

Healthy Brain Network. With the goal of collecting data from 10,000 children, adolescents, and young adults, HBN is another large study that will provide valuable insight into brain and behavior associations (Alexander et al. 2017). Relative to the ABCD study, HBN has some advantages along with a few disadvantages. A major advantage concerns recruitment. Through advertisements, HBN seeks families who have concerns about psychiatric symptoms in their child, and the study team offers comprehensive diagnostic evaluations along with recommendations for treatment when relevant. The expected result is that the sample will be enriched with many young people with a psychiatric diagnosis and elevated symptoms. Another advantage is that participants

undergo a clinician-administered K-SADS along with clinical consensus, which is thought to yield a more accurate diagnosis (Zelenina et al. 2023). In addition, with people between the ages of 5 and 21 being recruited, the project has a broad age range, making it possible to cross-sectionally examine how clinical conditions relate to brain function at different ages of development. Finally, because the catchment area is New York City, the sample is expected to be racially, ethnically, and economically diverse. As for disadvantages, it appears that data will be collected at only one time point, so longitudinal analysis will not be possible. Also, HBN is not a population-based sample, and so generalizability is an issue. An additional disadvantage is that since the social environment is not a focus of the work, measures of poverty are likely limited. Overall, because the large sample has high rates of psychopathology and a more reliable version of the K-SADS was used, HBN will provide an important contribution to the field.

Emerging data sets. Important new data sets will also become available soon including the HEALthy Brain and Child Development Study (HBCD), the Human Connectome Project—Development (HCP-D), the Study of Adolescent-to-Adult Neural Development (SAND), and ENIGMA-Environment. HBCD is analogous to the ABCD study, but it will span the first decade of life (Volkow et al. 2021). Moreover, the project will recruit people who used substances during pregnancy, and the 7,500 families in the sample from across the United States are expected to be diverse, including many participants who have been exposed to poverty. Therefore, given that data will be collected very early in development, often in the prenatal period, and data from multiple time points will be collected, this study will provide an incredibly rich resource for investigators to probe developmental associations among social environment, brain development, and mental health. HCP-D is expected to be composed of 1,300 healthy children, adolescents, and young adults (ages 5–21) from four cities, St. Louis, Boston, Los Angeles, and Minneapolis (Harms et al. 2018, Somerville et al. 2018). In addition, an adolescent subset is participating in a three-wave longitudinal component. Thus, developmental brain trajectories can be examined. Because much of the data are cross-sectional over a broad age range, the data will nicely complement the ABCD study. Moreover, the project is expected to have diversity with respect to SES, with ample representation of families with lower income.

SAND is recruiting participants from the Future of Families and Child Wellbeing Study (formerly called the Fragile Families and Child Wellbeing Study) (Reichman et al. 2001). Therefore, SAND has economic and socioemotional data at birth and at 1, 3, 5, 9, and 15 years (Hein et al. 2020). Moreover, although a full range of incomes is represented in the sample, there is substantial enrichment for low-income families, and, thus, it is well suited to examine correlates of poverty and development. The sample is also composed of two-thirds Black Americans, a population that is particularly underrepresented in research. The study is collecting mental health and MRI data from approximately 600 participants at two time points in young adulthood, with 237 having participated in adolescence for a total of three time points. Finally, within the ENIGMA Consortium framework (Novak et al. 2012), ENIGMA-Environment is bringing together a growing list of international contributors to examine potential effects of poverty-relevant variables, such as neighborhood economic markers, green spaces, and pollution levels, on brain structure and function (Ferschmann et al. 2022). Thus, along with the ABCD and HBN studies, these new data sets will provide important contributions to the field.

Randomized Controlled Trials

A common takeaway from recent work documenting small effect sizes for associations between brain and behavior is that the only path forward is to do large, consortia-based studies. However, in an excellent assessment of the field, Gratton and colleagues (2022) argue that the key to increasing

replicability is statistical power. One approach to increasing power while also separating cause and effect is to conduct randomized controlled trials (RCTs) with repeated measures before, during, and after the intervention or treatment. Of relevance to the present review, RCTs could be done using treatments or interventions, such as cash transfers, and in the context of environmental assessments, neuroimaging, and phenotyping. While there is enthusiasm in the field for conducting such RCTs, such enthusiasm must be tempered due to findings whereby the treatment conditions, including prekindergarten education and vouchers to move out of poverty, were linked to worse outcomes (Durkin et al. 2022, Kessler et al. 2014). These findings highlight the complexity of the issues and the need to have a more complete understanding of the relevant variables. In particular, these inconsistent results emphasize the need to gain a deeper comprehension of the underlying mechanisms that may explain the varying outcomes of poverty exposure.

Improved Sampling

Establishing representative samples is an important step in improving replicability. A sample or subset is representative if it reflects the population of interest. If the sample is not representative of a population (e.g., a convenience sample), it is unknown to whom the findings will generalize, and it may not be possible to recreate the sample to test for replication. Related to representativeness, we need greater consideration for the intersectionality of identities and experiences, including with minoritized groups such as individuals who have been exposed to interpersonal and structural racism. This requires representative samples that include more individuals who are historically underrepresented in neuroscience research (Falk et al. 2013, Garcini et al. 2022, Ricard et al. 2023). Several recommendations have been put forth to facilitate the recruitment of diverse samples. One is to improve community outreach and engagement in order to recruit underrepresented groups (Ocloo & Matthews 2016), and another is increased transparency in communication so that research could be disseminated in a way that would provide educational benefits for community members. To sustain research engagement, barriers to research participation should be reduced by diversifying the scientific community and improving cultural competence in research staff. These efforts would be beneficial for building trust between the scientific community and communities of color that were harmed in the past through participation in research. Moreover, this work is especially important for poverty research given that social class highly intersects with racial identity in the United States.

Although work using representative samples could facilitate generalizability and reproducibility, the goal of identifying robust and universal biomarkers of behavior may be more difficult to achieve in more diverse samples. This is because the issues concerning heterogeneity described above become even more complex in large, diverse samples. In particular, some neural markers may be idiosyncratic and person-specific depending on the multitude of complex interactions during the course of development between genetics and environmental conditions. Of note, an important recent study found that individuals who fall outside of predictable patterns (i.e., their stereotypical profile) are more likely to be misclassified using predictive models linking brain to behavior (Greene et al. 2022). This suggests that we need to be careful when drawing inferences about universal biomarkers that were obtained using large data sets because the environment, brain, and phenotypes are highly heterogeneous; resulting neural markers may not apply to a large number of people who defy normative patterns. Thus, in light of this limitation, it is important not only to identify to whom the findings generalize but also to determine to whom the findings do not generalize. This is especially important if we are interested in identifying markers of resilience to adversity because resilience in essence indicates a deviation from expected outcomes.

A final and crucial consideration for improved sampling is the inclusion of well-sampled longitudinal studies that allow for the analysis of developmental variables and trajectories. Given the

importance of the developmental timing of poverty experience as well as individual differences in how the brain develops, longitudinal data that provide information about poverty exposure and brain measures at multiple time points would allow for a better understanding of their interactions as well as increase reliability. Furthermore, cross-sectional data preclude inferences about causal associations and within-person change over time, which are important in identifying the long-term effects of poverty on brain and behavior.

Examining Resilience

In addition to examining risk, it is vital that we take a holistic approach and also look at resilience to poverty (Arredondo 2021, Arredondo et al. 2022). Within a dynamic system framework, resilience is defined as the ability to adapt to challenges that threaten the functioning, survival, and/or development of the organism (Masten et al. 2021). Fifty years of research on the topic has identified an array of factors across multiple systems that confer resilience, including sensitive caregiving, social support, sense of belonging, and optimism (Masten et al. 2021). By ignoring resilience and the socioeconomic variables that contribute to it, we not only do a disservice to the communities we are studying by losing the possibility of identifying routes for prevention and intervention strategies but also likely reduce reproducibility, since these variables are not accounted for in the model.

Replication Testing

In recent years, there has been a call to conduct out-of-sample or external replication testing (Gratton et al. 2022). Clearly, if findings are confirmed in another sample, faith is increased in the reliability of a particular finding, which moves us toward a more reproducible science. However, since demographic indices, environmental conditions, and other unaccounted-for variables may interact and impact brain development and mental health, it is a challenge to identify two or more samples that are comparable for replication testing. Vast differences in demographics and age ranges would make replication less likely. Nevertheless, if findings across samples do replicate, it provides powerful evidence for the finding. At the same time, if a replication fails, it may not be the case that the original finding was wrong; rather, the finding may accurately reflect conditions in one sample but not the other. Relatedly, borrowing from established procedures in epidemiology, neuroimaging researchers can implement a weighting approach to adjust for demographic differences between samples (Gard et al. 2023, Heeringa et al. 2017) to reduce sampling bias and increase generalizability of neural markers. One can also evaluate which demographic variables between samples may have prevented the findings from generalizing from one sample to the other.

New methods for external replication, such as meta-matching, have shown promise in leveraging neural network models that were trained using large samples that could be applied to behavioral prediction in smaller samples (Chopra et al. 2022, He et al. 2022). Recently, meta-matching was trained on the UK Biobank and then applied to small clinical samples where large data sets are not possible (Chopra et al. 2022, He et al. 2022). It was found that meta-matching increased the prediction of brain connectivity to cognitive function. Investigators may wish to be cautious in this approach as it makes the large data sets the ground truth. Nevertheless, it does offer a way for small studies to establish greater replicability between brain and behavior associations.

When the features of a sample are unique, and it does not seem reasonable to expect that the findings can replicate in another sample, within-sample replication may be the best option. There are two ways to accomplish this. First, split-half reliability can be done by subdividing data sets in two (Hardi et al. 2023a, Pronk et al. 2022). Second, one may randomly draw subsamples of

participants from the sample to assess the stability of the original findings (Hardi et al. 2023a, Tomasi & Volkow 2021). Thus, while not as optimal as out-of-sample replication, within-sample replication allows for the assessment of replicability in samples that are more distinct.

Increasing Interpretability

As we work toward developing reproducible links among poverty, brain, and behavior, one important challenge is to continue to center our research on benefiting individuals and families who are living in poverty. To this end, it is important that the discovery of neural markers of poverty and mental health disorders has relevance for interventions and policy. When work on developmental human neuroscience began, it often used neuroscience-driven hypotheses from animal models to facilitate interpretation and provide a bridge to applications. However, as sample sizes become larger in the era of big data, and neural models become increasingly complex, there are trade-offs between model complexity and interpretability. For instance, increasing the number of features in machine learning approaches can boost predictive ability but result in an increase in model complexity that makes it more difficult to precisely identify targets for interventions or to test theory-driven hypotheses (Nielsen et al. 2020, Tejavibulya et al. 2022). Thus, as the field progresses to adopt these novel methodological approaches, striking a balance between complexity and interpretability is important. In particular, we suggest that the field integrate data science approaches with hypotheses from animal models and interventions, such as cash transfer RCTs. The integration of these three elements will increase reliability and replicability while also helping to ground findings in mechanisms and applications.

CONCLUSION

Despite the rapid expansion of the field linking brain measures to poverty and mental health, concerns about replicability remain. Some compelling evidence indicates that poverty has replicable links to brain development, but associations between brain and mental health are less consistent. Heterogeneity in brain, mental health symptoms, and links between the two may contribute to this inconsistency. The incorporation of large data sets can increase replicability, but they, too, have limitations. Moving forward, we suggest that integrating data science, neuroscience models, and intervention research is essential for facilitating greater replicability, interpretability, and links to policy.

DISCLOSURE STATEMENT

The authors are not aware of any affiliations, memberships, funding, or financial holdings that might be perceived as affecting the objectivity of this review.

ACKNOWLEDGMENTS

This work was supported by the National Institute of Mental Health award R01 MH121079 (principal investigators Hyde, Mitchell, and Monk) and the Eunice Kennedy Shriver National Institute of Child Health and Human Development award T32 HD007109 (principal investigators Gelman and Monk).

LITERATURE CITED

Akee RKQ, Copeland WE, Keeler G, Angold A, Costello EJ. 2010. Parents' incomes and children's outcomes: a quasi-experiment using transfer payments from casino profits. *Am. Econ. J. Appl. Econ.* 2(1):86–115

Alexander LM, Escalera J, Ai L, Andreotti C, Febre K, et al. 2017. An open resource for transdiagnostic research in pediatric mental health and learning disorders. *Sci. Data* 4(1):170181

Arredondo MM. 2021. Shining a light on cultural neuroscience: recommendations on the use of fNIRS to study how sociocultural contexts shape the brain. *Cult. Divers. Ethn. Minor. Psychol.* 29(1):106–17

Arredondo MM, Garcini LM, McLaughlin KA. 2022. Integration of equity and diversity frameworks to advance biological psychiatry. *Biol. Psychiatry Cogn. Neurosci. Neuroimaging* 7(12):1198–99

Axinn WG, Choi KW, Ghimire DJ, Cole F, Hermosilla S, et al. 2022. Community-level social support infrastructure and adult onset of major depressive disorder in a South Asian postconflict setting. *JAMA Psychiatry* 79(3):243–49

Bandettini PA, Gonzalez-Castillo J, Handwerker D, Taylor P, Chen G, et al. 2022. The challenge of BWAs: unknown unknowns in feature space and variance. *Med* 3(8):526–31

Birn RM, Molloy EK, Patriat R, Parker T, Meier TB, et al. 2013. The effect of scan length on the reliability of resting-state fMRI connectivity estimates. *NeuroImage* 83:550–58

Biswal BB, Mennes M, Zuo X-N, Gohel S, Kelly C, et al. 2010. Toward discovery science of human brain function. *PNAS* 107(10):4734–39

Biswal BB, Yetkin FZ, Haughton VM, Hyde JS. 1995. Functional connectivity in the motor cortex of resting human brain using echo-planar MRI. *Magn. Reson. Med.* 34(4):537–41

Bloom PA, VanTieghem M, Gabard-Durnam L, Gee DG, Flannery J, et al. 2022. Age-related change in task-evoked amygdala—prefrontal circuitry: a multiverse approach with an accelerated longitudinal cohort aged 4–22 years. *Hum. Brain Mapp.* 43(10):3221–44

Boyce WT, Ellis BJ. 2005. Biological sensitivity to context: I. An evolutionary-developmental theory of the origins and functions of stress reactivity. *Dev. Psychopathol.* 17(2):271–301

Brito NH, Noble KG. 2014. Socioeconomic status and structural brain development. *Front. Neurosci.* 8:276

Brooks-Gunn J, Duncan GJ. 1997. The effects of poverty on children. *Future Child.* 7(2):55–71

Button KS, Ioannidis JPA, Mokrysz C, Nosek BA, Flint J, et al. 2013. Power failure: why small sample size undermines the reliability of neuroscience. *Nat. Rev. Neurosci.* 14(5):365–76

Bzdok D, Meyer-Lindenberg A. 2018. Machine learning for precision psychiatry: opportunities and challenges. *Biol. Psychiatry Cogn. Neurosci. Neuroimaging* 3(3):223–30

Carp J. 2013. Better living through transparency: improving the reproducibility of fMRI results through comprehensive methods reporting. *Cogn. Affect. Behav. Neurosci.* 13(3):660–66

Casey BJ, Cannonier T, Conley MI, Cohen AO, Barch DM, et al. 2018. The Adolescent Brain Cognitive Development (ABCD) study: imaging acquisition across 21 sites. *Dev. Cogn. Neurosci.* 32:43–54

Chen G, Pine DS, Brotman MA, Smith AR, Cox RW, et al. 2021. Trial and error: a hierarchical modeling approach to test-retest reliability. *NeuroImage* 245:118647

Chen J, Tam A, Kebets V, Orban C, Ooi LQR, et al. 2022. Shared and unique brain network features predict cognitive, personality, and mental health scores in the ABCD study. *Nat Commun.* 13(1):2217

Chetty R, Hendren N, Katz LF. 2016. The effects of exposure to better neighborhoods on children: new evidence from the Moving to Opportunity experiment. *Am. Econ. Rev.* 106(4):855–902

Chopra S, Dhamala E, Lawhead C, Ricard JA, Orchard ER, et al. 2022. Reliable and generalizable brain-based predictions of cognitive functioning across common psychiatric illness. medRxiv 2228323. **https://doi.org/10.1101/2022.12.08.22283232**

Cicchetti D, Rogosch FA. 1996. Equifinality and multifinality in developmental psychopathology. *Dev. Psychopathol.* 8(4):597–600

Cohen AD, Yang B, Fernandez B, Banerjee S, Wang Y. 2021. Improved resting state functional connectivity sensitivity and reproducibility using a multiband multi-echo acquisition. *NeuroImage* 225:117461

Conger RD, Wallace LE, Sun Y, Simons RL, McLoyd VC, et al. 2002. Economic pressure in African American families: a replication and extension of the family stress model. *Dev. Psychol.* 38(2):179–93

Cosgrove KT, McDermott TJ, White EJ, Mosconi MW, Thompson WK, et al. 2022. Limits to the generalizability of resting-state functional magnetic resonance imaging studies of youth: an examination of ABCD Study® baseline data. *Brain Imaging Behav.* 16:1919–25

Costello EJ, Compton SN, Keeler G, Angold A. 2003. Relationships between poverty and psychopathology: a natural experiment. *JAMA* 290(15):2023–29

Dahl GB, Lochner L. 2012. The impact of family income on child achievement: evidence from the Earned Income Tax Credit. *Am. Econ. Rev.* 102(5):1927–56

Dhollander T, Clemente A, Singh M, Boonstra F, Civier O, et al. 2021. Fixel-based analysis of diffusion MRI: methods, applications, challenges and opportunities. *NeuroImage* 241:118417

Di Martino A, Yan C-G, Li Q, Denio E, Castellanos FX, et al. 2014. The autism brain imaging data exchange: towards large-scale evaluation of the intrinsic brain architecture in autism. *Mol. Psychiatry* 19(6):659–67

Drysdale AT, Grosenick L, Downar J, Dunlop K, Mansouri F, et al. 2017. Resting-state connectivity biomarkers define neurophysiological subtypes of depression. *Nat. Med.* 23(1):28–38

Duncan GJ, Magnuson KA. 2003. Off with Hollingshead: socioeconomic resources, parenting, and child development. In *Socioeconomic Status, Parenting, and Child Development*, ed. MH Bornstein, RH Bradley, pp. 83–106. Mahwah, NJ: Lawrence Erlbaum Assoc.

Duncan GJ, Magnuson K, Votruba-Drzal E. 2014. Boosting family income to promote child development. *Future Child.* 24(1):99–120

Duncan GJ, Morris PA, Rodrigues C. 2011. Does money really matter? Estimating impacts of family income on young children's achievement with data from random-assignment experiments. *Dev. Psychol.* 47:1263–79

Durkin K, Lipsey MW, Dale CF, Wiesen SE. 2022. Effects of a statewide pre-kindergarten program on children's achievement and behavior through sixth grade. *Dev. Psychol.* 58(3):470–84

Durstewitz D, Koppe G, Meyer-Lindenberg A. 2019. Deep neural networks in psychiatry. *Mol. Psychiatry* 24(11):1583–98

Elliott ML, Knodt AR, Cooke M, Kim MJ, Melzer TR, et al. 2019. General functional connectivity: shared features of resting-state and task fMRI drive reliable and heritable individual differences in functional brain networks. *NeuroImage* 189:516–32

Elliott ML, Knodt AR, Hariri AR. 2021. Striving toward translation: strategies for reliable fMRI measurement. *Trends Cogn. Sci.* 25(9):776–87

Elliott ML, Knodt AR, Ireland D, Morris ML, Poulton R, et al. 2020. What is the test-retest reliability of common task-functional MRI measures? New empirical evidence and a meta-analysis. *Psychol. Sci.* 31(7):792–806

Ellwood-Lowe ME, Whitfield-Gabrieli S, Bunge SA. 2021. Brain network coupling associated with cognitive performance varies as a function of a child's environment in the ABCD study. *Nat. Commun.* 12(1):7183

Evans GW. 2004. The environment of childhood poverty. *Am. Psychol.* 59(2):77–92

Fair DA, Dosenbach NUF, Moore AH, Satterthwaite TD, Milham MP. 2021. Developmental cognitive neuroscience in the era of networks and big data: strengths, weaknesses, opportunities, and threats. *Annu. Rev. Dev. Psychol.* 3:249–75

Falk EB, Hyde LW, Mitchell C, Faul J, Gonzalez R, et al. 2013. What is a representative brain? Neuroscience meets population science. *PNAS* 110(44):17615–22

Feczko E, Miranda-Dominguez O, Marr M, Graham AM, Nigg JT, et al. 2019. The heterogeneity problem: approaches to identify psychiatric subtypes. *Trends Cogn. Sci.* 23(7):584–601

Feilong M, Guntupalli JS, Haxby JV. 2021. The neural basis of intelligence in fine-grained cortical topographies. *eLife* 10:e64058

Ferschmann L, Bos MGN, Herting MM, Mills KL, Tamnes CK. 2022. Contextualizing adolescent structural brain development: environmental determinants and mental health outcomes. *Curr. Opin. Psychol.* 44:170–76

Finn ES. 2021. Is it time to put rest to rest? *Trends Cogn. Sci.* 25(12):1021–32

Finn ES, Bandettini PA. 2021. Movie-watching outperforms rest for functional connectivity-based prediction of behavior. *NeuroImage* 235:117963

Finn ES, Shen X, Scheinost D, Rosenberg MD, Huang J, et al. 2015. Functional connectome fingerprinting: identifying individuals using patterns of brain connectivity. *Nat. Neurosci.* 18(11):1664–71

Fox MD, Snyder AZ, Vincent JL, Corbetta M, Van Essen DC, et al. 2005. The human brain is intrinsically organized into dynamic, anticorrelated functional networks. *PNAS* 102(27):9673–78

Garavan H, Bartsch H, Conway K, Decastro A, Goldstein RZ, et al. 2018. Recruiting the ABCD sample: design considerations and procedures. *Dev. Cogn. Neurosci.* 32:16–22

Garcini LM, Arredondo MM, Berry O, Church JA, Fryberg S, et al. 2022. Increasing diversity in developmental cognitive neuroscience: a roadmap for increasing representation in pediatric neuroimaging research. *Dev. Cogn. Neurosci.* 58:101167

Gard AM, Hyde LW, Heeringa SG, West BT, Mitchell C. 2023. Why weight? Analytic approaches for large-scale population neuroscience data. *Dev. Cogn. Neurosci.* 59:101196

Garmezy N. 1993. Children in poverty: resilience despite risk. *Psychiatry* 56(1):127–36

Gershoff ET, Aber JL, Raver CC, Lennon MC. 2007. Income is not enough: incorporating material hardship into models of income associations with parenting and child development. *Child Dev.* 78(1):70–95

Giedd JN, Blumenthal J, Jeffries NO, Castellanos FX, Liu H, et al. 1999. Brain development during childhood and adolescence: a longitudinal MRI study. *Nat. Neurosci.* 2(10):861–63

Goetschius LG, Hein TC, McLanahan SS, Brooks-Gunn J, McLoyd VC, et al. 2020a. Association of childhood violence exposure with adolescent neural network density. *JAMA Netw. Open* 3(9):e2017850

Goetschius LG, Hein TC, Mitchell C, Lopez-Duran NL, McLoyd VC, et al. 2020b. Childhood violence exposure and social deprivation predict adolescent amygdala-orbitofrontal cortex white matter connectivity. *Dev. Cogn. Neurosci.* 45:100849

Goodlad JK, Marcus DK, Fulton JJ. 2013. Lead and attention-deficit/hyperactivity disorder (ADHD) symptoms: a meta-analysis. *Clin. Psychol. Rev.* 33(3):417–25

Gordon EM, Laumann TO, Gilmore AW, Newbold DJ, Greene DJ, et al. 2017. Precision functional mapping of individual human brains. *Neuron* 95(4):791–807.e7

Gorgolewski KJ, Auer T, Calhoun VD, Craddock RC, Das S, et al. 2016. The brain imaging data structure, a format for organizing and describing outputs of neuroimaging experiments. *Sci. Data* 3(1):160044

Gratton C, Kraus BT, Greene DJ, Gordon EM, Laumann TO, et al. 2020. Defining individual-specific functional neuroanatomy for precision psychiatry. *Biol. Psychiatry* 88(1):28–39

Gratton C, Laumann TO, Nielsen AN, Greene DJ, Gordon EM, et al. 2018. Functional brain networks are dominated by stable group and individual factors, not cognitive or daily variation. *Neuron* 98(2):439–52.e5

Gratton C, Nelson SM, Gordon EM. 2022. Brain-behavior correlations: two paths toward reliability. *Neuron* 110(9):1446–49

Greene AS, Gao S, Scheinost D, Constable RT. 2018. Task-induced brain state manipulation improves prediction of individual traits. *Nat. Commun.* 9(1):2807

Greene AS, Shen X, Noble S, Horien C, Hahn CA, et al. 2022. Brain-phenotype models fail for individuals who defy sample stereotypes. *Nature* 609(7925):109–18

Hackman DA, Farah MJ. 2009. Socioeconomic status and the developing brain. *Trends Cogn. Sci.* 13(2):65–73

Hair NL, Hanson JL, Wolfe BL, Pollak SD. 2015. Association of child poverty, brain development, and academic achievement. *JAMA Pediatr.* 169(9):822–29

Hanson JL, Hair N, Shen DG, Shi F, Gilmore JH, et al. 2013. Family poverty affects the rate of human infant brain growth. *PLOS ONE* 8(12):e80954

Hardi FA, Goetschius LG, McLoyd V, Lopez-Duran NL, Mitchell C, et al. 2023a. Adolescent functional network connectivity prospectively predicts adult anxiety symptoms related to perceived COVID-19 economic adversity. *J. Child Psychol. Psychiatry* 64(6):918–29

Hardi FA, Goetschius LG, Tillem S, McLoyd V, Brooks-Gunn J, et al. 2023b. Early childhood household instability, adolescent structural neural network architecture, and young adulthood depression: a 21-year longitudinal study. *Dev. Cogn. Neurosci.* 61:101253

Hardi FA, Goetschius LG, Peckins MK, Brooks-Gunn J, McLanahan SS, et al. 2022. Differential developmental associations of material hardship exposure and adolescent amygdala–prefrontal cortex white matter connectivity. *J. Cogn. Neurosci.* 34(10):1866–91

Hardoon DR, Szedmak S, Shawe-Taylor J. 2004. Canonical correlation analysis: an overview with application to learning methods. *Neural Comput.* 16(12):2639–64

Hardt J, Rutter M. 2004. Validity of adult retrospective reports of adverse childhood experiences: review of the evidence. *J. Child Psychol. Psychiatry* 45(2):260–73

Harms MP, Somerville LH, Ances BM, Andersson J, Barch DM, et al. 2018. Extending the Human Connectome Project across ages: imaging protocols for the Lifespan Development and Aging projects. *NeuroImage* 183:972–84

Haxby JV, Guntupalli JS, Connolly AC, Halchenko YO, Conroy BR, et al. 2011. A common, high-dimensional model of the representational space in human ventral temporal cortex. *Neuron* 72(2):404–16

He T, An L, Chen P, Chen J, Feng J, et al. 2022. Meta-matching as a simple framework to translate phenotypic predictive models from big to small data. *Nat. Neurosci.* 25(6):795–804

Heeringa SG, West BT, Berglund PA. 2017. *Applied Survey Data Analysis*. Boca Raton, FL: CRC Press

Hein TC, Goetschius LG, McLoyd VC, Brooks-Gunn J, McLanahan SS, et al. 2020. Childhood violence exposure and social deprivation are linked to adolescent threat and reward neural function. *Soc. Cogn. Affect. Neurosci.* 15(11):1252–59

Hensch TK, Bilimoria PM. 2012. Re-opening windows: manipulating critical periods for brain development. *Cerebrum* 2012:11

Jensen SKG, Berens AE, Nelson CA. 2017. Effects of poverty on interacting biological systems underlying child development. *Lancet Child Adolesc. Health* 1(3):225–39

Jiang R, Calhoun VD, Cui Y, Qi S, Zhuo C, et al. 2020. Multimodal data revealed different neurobiological correlates of intelligence between males and females. *Brain Imaging Behav.* 14(5):1979–93

Kaufman J, Birmaher B, Brent D, Rao U, Flynn C, et al. 1997. Schedule for Affective Disorders and Schizophrenia for School-Age Children-Present and Lifetime Version (K-SADS-PL): initial reliability and validity data. *J. Am. Acad. Child Adolesc. Psychiatry* 36(7):980–88

Kessler RC, Duncan GJ, Gennetian LA, Katz LF, Kling JR, et al. 2014. Associations of housing mobility interventions for children in high-poverty neighborhoods with subsequent mental disorders during adolescence. *JAMA* 311(9):937–47

Keyes KM, Gary D, O'Malley PM, Hamilton A, Schulenberg J. 2019. Recent increases in depressive symptoms among US adolescents: trends from 1991 to 2018. *Soc. Psychiatry Psychiatr. Epidemiol.* 54(8):987–96

Kragel PA, Han X, Kraynak TE, Gianaros PJ, Wager TD. 2021. Functional MRI can be highly reliable, but it depends on what you measure: a commentary on Elliott et al. 2020. *Psychol. Sci.* 32(4):622–26

Kundu P, Brenowitz ND, Voon V, Worbe Y, Vértes PE, et al. 2013. Integrated strategy for improving functional connectivity mapping using multiecho fMRI. *PNAS* 110(40):16187–92

Lazzarino AI, Hamer M, Stamatakis E, Steptoe A. 2013. The combined association of psychological distress and socioeconomic status with all-cause mortality: a national cohort study. *JAMA Intern. Med.* 173(1):22–27

Li X, Ai L, Giavasis S, Jin H, Feczko E, et al. 2022. Moving beyond processing and analysis-related variation in neuroscience. bioRxiv 470790. **https://doi.org/10.1101/2021.12.01.470790**

Lindsay DS. 2015. Replication in psychological science. *Psychol. Sci.* 26(12):1827–32

Lynch CJ, Power JD, Scult MA, Dubin M, Gunning FM, et al. 2020. Rapid precision functional mapping of individuals using multi-echo fMRI. *Cell Rep.* 33(12):108540

Månsson KNT, Waschke L, Manzouri A, Furmark T, Fischer H, et al. 2022. Moment-to-moment brain signal variability reliably predicts psychiatric treatment outcome. *Biol. Psychiatry* 91(7):658–66

Marek S, Tervo-Clemmens B, Calabro FJ, Montez DF, Kay BP, et al. 2022. Reproducible brain-wide association studies require thousands of individuals. *Nature* 603(7902):654–60

Marquand AF, Kia SM, Zabihi M, Wolfers T, Buitelaar JK, et al. 2019. Conceptualizing mental disorders as deviations from normative functioning. *Mol. Psychiatry* 24(10):1415–24

Masten AS, Lucke CM, Nelson KM, Stallworthy IC. 2021. Resilience in development and psychopathology: multisystem perspectives. *Annu. Rev. Clin. Psychol.* 17:521–49

McLaughlin KA, Sheridan MA. 2016. Beyond cumulative risk: a dimensional approach to childhood adversity. *Curr. Dir. Psychol. Sci.* 25(4):239–45

McLoyd VC. 1998. Socioeconomic disadvantage and child development. *Am. Psychol.* 20:185–204

McLoyd VC. 2019. How children and adolescents think about, make sense of, and respond to economic inequality: Why does it matter? *Dev. Psychol.* 55(3):592–600

Milne BJ, D'Souza S, Andersen SH, Richmond-Rakerd LS. 2022. Use of population-level administrative data in developmental science. *Annu. Rev. Dev. Psychol.* 4:447–68

Miranda-Dominguez O, Mills BD, Carpenter SD, Grant KA, Kroenke CD, et al. 2014. Connectotyping: model based fingerprinting of the functional connectome. *PLOS ONE* 9(11):e111048

Molenaar PCM. 2004. A manifesto on psychology as idiographic science: bringing the person back into scientific psychology, this time forever. *Meas. Interdiscip. Res. Perspect.* 2(4):201–18

Montag C, Quintana DS. 2023. Digital phenotyping in molecular psychiatry—a missed opportunity? *Mol. Psychiatry* 28(1):6–9

Mueller CW, Parcel TL. 1981. Measures of socioeconomic status: alternatives and recommendations. *Child Dev.* 52(1):13–30

Mueller S, Wang D, Fox MD, Yeo BTT, Sepulcre J, et al. 2013. Individual variability in functional connectivity architecture of the human brain. *Neuron* 77(3):586–95

Natl. Inst. Child Health Hum. Dev. Early Child Care Res. Netw. 2005. Duration and developmental timing of poverty and children's cognitive and social development from birth through third grade. *Child Dev.* 76(4):795–810

Nielsen AN, Barch DM, Petersen SE, Schlaggar BL, Greene DJ. 2020. Machine learning with neuroimaging: evaluating its applications in psychiatry. *Biol. Psychiatry Cogn. Neurosci. Neuroimaging* 5(8):791–98

Nikolaidis A, Chen AA, He X, Shinohara R, Vogelstein J, et al. 2022. Suboptimal phenotypic reliability impedes reproducible human neuroscience. bioRxiv 501193. **https://doi.org/10.1101/2022.07.22.501193**

Noble S, Spann MN, Tokoglu F, Shen X, Constable RT, et al. 2017. Influences on the test-retest reliability of functional connectivity MRI and its relationship with behavioral utility. *Cereb. Cortex* 27(11):5415–29

Novak NM, Stein JL, Medland SE, Hibar DP, Thompson PM, et al. 2012. EnigmaVis: online interactive visualization of genome-wide association studies of the Enhancing NeuroImaging Genetics through Meta-Analysis (ENIGMA) consortium. *Twin Res. Hum. Genet.* 15(3):414–18

Ocloo J, Matthews R. 2016. From tokenism to empowerment: progressing patient and public involvement in healthcare improvement. *BMJ Qual. Saf.* 25(8):626–32

Ooi LQR, Chen J, Shaoshi Z, Kong R, Tam A, et al. 2022. Comparison of individualized behavioral predictions across anatomical, diffusion and functional connectivity MRI. *NeuroImage* 263:119636

Orrù G, Pettersson-Yeo W, Marquand AF, Sartori G, Mechelli A. 2012. Using Support Vector Machine to identify imaging biomarkers of neurological and psychiatric disease: a critical review. *Neurosci. Biobehav. Rev.* 36(4):1140–52

Parkes L, Fulcher B, Yücel M, Fornito A. 2018. An evaluation of the efficacy, reliability, and sensitivity of motion correction strategies for resting-state functional MRI. *NeuroImage* 171:415–36

Poldrack RA, Baker CI, Durnez J, Gorgolewski KJ, Matthews PM, et al. 2017. Scanning the horizon: towards transparent and reproducible neuroimaging research. *Nat. Rev. Neurosci.* 18(2):115–26

Poser BA, Versluis MJ, Hoogduin JM, Norris DG. 2006. BOLD contrast sensitivity enhancement and artifact reduction with multiecho EPI: parallel-acquired inhomogeneity-desensitized fMRI. *Magn. Reson. Med.* 55(6):1227–35

Price RB, Lane S, Gates K, Kraynak TE, Horner MS, et al. 2017. Parsing heterogeneity in the brain connectivity of depressed and healthy adults during positive mood. *Biol. Psychiatry* 81(4):347–57

Pronk T, Molenaar D, Wiers RW, Murre J. 2022. Methods to split cognitive task data for estimating split-half reliability: a comprehensive review and systematic assessment. *Psychon. Bull. Rev.* 29(1):44–54

Rakesh D, Whittle S. 2021. Socioeconomic status and the developing brain—a systematic review of neuroimaging findings in youth. *Neurosci. Biobehav. Rev.* 130:379–407

Regier DA, Narrow WE, Clarke DE, Kraemer HC, Kuramoto SJ, et al. 2013. DSM-5 field trials in the United States and Canada, part II: test-retest reliability of selected categorical diagnoses. *Am. J. Psychiatry* 170(1):59–70

Reichman NE, Teitler JO, Garfinkel I, McLanahan SS. 2001. Fragile Families: sample and design. *Child. Youth. Serv. Rev.* 23(4):303–26

Ricard JA, Parker TC, Dhamala E, Kwasa J, Allsop A, et al. 2023. Confronting racially exclusionary practices in the acquisition and analyses of neuroimaging data. *Nat. Neurosci.* 26(1):4–11

Rutherford S. 2020. The promise of machine learning for psychiatry. *Biol. Psychiatry* 88(11):E53–55

Rutherford S, Kia SM, Wolfers T, Fraza C, Zabihi M, et al. 2022. The normative modeling framework for computational psychiatry. *Nat. Protoc.* 17(7):1711–34

Saez E, Zucman G. 2014. Wealth inequality in the United States since 1913: evidence from capitalized income tax data. *Q. J. Econ.* 131(2):519–78

Satterthwaite TD, Elliott MA, Ruparel K, Loughead J, Prabhakaran K, et al. 2014. Neuroimaging of the Philadelphia Neurodevelopmental Cohort. *NeuroImage* 86:544–53

Schmaal L. 2022. The search for clinically useful neuroimaging markers of depression—a worthwhile pursuit or a futile quest? *JAMA Psychiatry* 79(9):845–46

Schulz M-A, Bzdok D, Haufe S, Haynes J-D, Ritter K. 2022. Performance reserves in brain-imaging-based phenotype prediction. bioRxiv 470790. **https://doi.org/10.1101/2021.12.01.470790**

Seider NA, Adeyemo B, Miller R, Newbold DJ, Hampton JM, et al. 2022. Accuracy and reliability of diffusion imaging models. *NeuroImage* 254:119138

Shankman SA, Funkhouser CJ, Klein DN, Davila J, Lerner D, et al. 2018. Reliability and validity of severity dimensions of psychopathology assessed using the Structured Clinical Interview for DSM-5 (SCID). *Int. J. Methods Psychiatr. Res.* 27(1):e1590

Shavers VL. 2007. Measurement of socioeconomic status in health disparities research. *J. Natl. Med. Assoc.* 99(9):1013–23

Somerville LH, Bookheimer SY, Buckner RL, Burgess GC, Curtiss SW, et al. 2018. The Lifespan Human Connectome Project in Development: a large-scale study of brain connectivity development in 5–21 year olds. *NeuroImage* 183:456–68

Spisak T, Bingel U, Wager T. 2022. Replicable multivariate BWAS with moderate sample sizes**.** bioRxiv 497072. **https://doi.org/10.1101/2022.06.22.497072**

Sui J, Jiang R, Bustillo J, Calhoun V. 2020. Neuroimaging-based individualized prediction of cognition and behavior for mental disorders and health: methods and promises. *Biol. Psychiatry* 88(11):818–28

Sui J, Qi S, van Erp TGM, Bustillo J, Jiang R, et al. 2018. Multimodal neuromarkers in schizophrenia via cognition-guided MRI fusion. *Nat. Commun.* 9(1):3028

Taylor RL, Cooper SR, Jackson JJ, Barch DM. 2020. Assessment of neighborhood poverty, cognitive function, and prefrontal and hippocampal volumes in children. *JAMA Netw. Open* 3(11):e2023774

Tejavibulya L, Rolison M, Gao S, Liang Q, Peterson H, et al. 2022. Predicting the future of neuroimaging predictive models in mental health. *Mol. Psychiatry* 27(8):3129–37

Tiego J, Martin EA, DeYoung CG, Hagan K, Cooper SE, et al. 2023. Precision behavioral phenotyping as a strategy for uncovering the biological correlates of psychopathology. *Nat. Mental Health* 1:304–15

Tomasi D, Volkow ND. 2021. Associations of family income with cognition and brain structure in USA children: prevention implications. *Mol. Psychiatry* 26(11):6619–29

Tooley UA, Bassett DS, Mackey AP. 2021. Environmental influences on the pace of brain development. *Nat. Rev. Neurosci.* 22(6):372–84

Tournier J-D, Calamante F, Connelly A. 2012. MRtrix: diffusion tractography in crossing fiber regions. *Int. J. Imaging Syst. Technol.* 22(1):53–66

Townsend L, Kobak K, Kearney C, Milham M, Andreotti C, et al. 2020. Development of three web-based computerized versions of the Kiddie Schedule for Affective Disorders and Schizophrenia child psychiatric diagnostic interview: preliminary validity data. *J. Am. Acad. Child Adolesc. Psychiatry* 59(2):309–25

Troller-Renfree SV, Costanzo MA, Duncan GJ, Magnuson K, Gennetian LA, et al. 2022. The impact of a poverty reduction intervention on infant brain activity. *PNAS* 119(5):e2115649119

Viruell-Fuentes EA, Miranda PY, Abdulrahim S. 2012. More than culture: structural racism, intersectionality theory, and immigrant health. *Soc. Sci. Med.* 75(12):2099–106

Volkow ND, Gordon JA, Freund MP. 2021. The Healthy Brain and Child Development Study—shedding light on opioid exposure, COVID-19, and health disparities. *JAMA Psychiatry* 78(5):471

Wang H-T, Smallwood J, Mourao-Miranda J, Xia CH, Satterthwaite TD, et al. 2020. Finding the needle in a high-dimensional haystack: canonical correlation analysis for neuroscientists. *NeuroImage* 216:116745

Winter NR, Leenings R, Ernsting J, Sarink K, Fisch L, et al. 2022. Quantifying deviations of brain structure and function in major depressive disorder across neuroimaging modalities. *JAMA Psychiatry* 79(9):879–88

Xia CH, Ma Z, Ciric R, Gu S, Betzel RF, et al. 2018. Linked dimensions of psychopathology and connectivity in functional brain networks. *Nat. Commun.* 9(1):3003

Zelenina M, Pine D, Stringaris A, Nielson D. 2023. Validation of CBCL depression scores of adolescents in three independent datasets. *Dev. Cogn. Neurosci.* In press. **https://doi.org/10.31234/osf.io/z956k**