Review

Biological Psychiatry

Toward Neurosubtypes in Autism

Seok-Jun Hong, Joshua T. Vogelstein, Alessandro Gozzi, Boris C. Bernhardt, B.T. Thomas Yeo, Michael P. Milham, and Adriana Di Martino

ABSTRACT

There is a consensus that substantial heterogeneity underlies the neurobiology of autism spectrum disorder (ASD). As such, it has become increasingly clear that a dissection of variation at the molecular, cellular, and brain network domains is a prerequisite for identifying biomarkers. Neuroimaging has been widely used to characterize atypical brain patterns in ASD, although findings have varied across studies. This is due, at least in part, to a failure to account for neurobiological heterogeneity. Here, we summarize emerging data-driven efforts to delineate more homogeneous ASD subgroups at the level of brain structure and function—that is, neurosubtyping. We break this pursuit into key methodological steps: the selection of diagnostic samples, neuroimaging features, algorithms, and validation approaches. Although preliminary and methodologically diverse, current studies generally agree that at least 2 to 4 distinct ASD neurosubtypes may exist. Their identification improved symptom prediction and diagnostic label accuracy above and beyond group average comparisons. Yet, this nascent literature has shed light onto challenges and gaps. These include 1) the need for wider and more deeply transdiagnostic samples collected while minimizing artifacts (e.g., head motion), 2) quantitative and unbiased methods for feature selection and multimodal fusion, 3) greater emphasis on algorithms' ability to capture hybrid dimensional and categorical models of ASD, and 4) systematic independent replications and validations that integrate different units of analyses across multiple scales. Solutions aimed to address these challenges and gaps are discussed for future avenues leading toward a comprehensive understanding of the mechanisms underlying ASD heterogeneity.

Keywords: Autism, Bayesian modeling, Data-driven clustering, Neuroimaging, Replicability, Subtyping

https://doi.org/10.1016/j.biopsych.2020.03.022

Clinical and etiological heterogeneity remain major obstacles to biomarker identification in psychiatry. Recognizing the limitations of current nosology (1–3), the past decade has witnessed pressing calls to partition psychiatric heterogeneity into more homogeneous groups—that is, subtyping (4–8). Here, we spotlight autism spectrum disorder (ASD), which with its early-childhood onset, remarkable clinical and biological variability (9–12), and increasing prevalence (13) epitomizes psychiatry's needs and challenges in subtyping.

Unlike other psychiatric disorders characterized by a spectrum of symptom severity and shared variance [e.g., schizophrenia, mood, anxiety (14,15)], individual variation in ASD goes beyond core diagnostic symptoms (i.e., impairment in social communicative skills and restricted/repetitive behaviors/interests) and associated psychopathology. ASD is often complicated by accompanying heterogeneity in the type of onset (16), language skills (17), intellectual abilities (18), and medical conditions (19). While high variability in the behavioral presentation of ASD has been recognized since its early descriptions (20,21), increases in prevalence over the past 40 years [from 0.05% in 1966 to ~2% in 2019 (22,23)] have likely further unveiled its heterogeneity. Evidence that ASD heterogeneity is increasingly detected comes from a recent meta-analysis showing that effect sizes for diagnostic group mean differences in ASD have decreased over time (24). The uniqueness of ASD is further accentuated by its onset during the first years of life combined with a long-term course. This is substantiated by dramatic variation in developmental trajectories (17,25–29), resulting functional outcomes (30), and response to treatment (31)—all of which underscore the need for a strong developmental perspective in subtyping approaches. The complexity of the ASD picture has also been highlighted by advances in genomics and neurobiology studies, which have collectively pointed toward multiple etiological pathways (9,12,32,33). The resulting range of clinical and biological ASD presentations represents both a formidable challenge and motivation for subtyping.

Subtyping approaches can be characterized in terms of both the unit(s) of analysis on which variation among individuals is indexed and the nature of the algorithm(s) used to sort individuals into groups (Supplemental Figure S1). Target units of analyses can be categorized under behavior and biology (34). The behavior-based units are thought to indirectly index variation of underlying biology (35) and vice versa. Regardless of the unit(s) of analyses, quantitative subtyping algorithm can be broadly categorized as supervised (i.e., label driven), unsupervised (i.e., data driven), or their hybrids. These approaches leverage univariate or multivariate statistics, each having advantages and disadvantages, as detailed elsewhere (36–38).

ISSN: 0006-3223

In ASD, until recently, behavior has been the predominant unit of analysis, whether using theoretically motivated or data-driven approaches (4,39). While the results of these efforts have captured relevant clinical aspects of autism, their findings have been neither exhaustive nor reliably distinguishable (40). As such, the DSM-5 (41) and the newest version of the ICD (42) have retreated to a single diagnostic category of ASD. This underscores 1) the need for more detailed behavioral phenotyping (11) and 2) the substantial gap in identifying subtypes anchored to biological features relevant to etiology and/or pathophysiology. In other fields of medicine, the detection of biologically based subtypes has led to earlier and more accurate recognition, treatment selection, and outcome monitoring [e.g., congestive heart failure (43)].

The differing pathophysiology and etiologies of ASD remain to be fully elucidated. However, owing to the pivotal contribution of genetic factors to ASD risk [heritability: 0.69–0.83 (44)], genomic analyses have led endeavors aimed to identify ASD underlying etiology and/or pathophysiology [detailed in (9,12,32,33)]. A number of studies examined the contribution of genetic ASD etiology yielding tremendous progress, particularly during the past decade. Results indicated that both common polygenic variation and rare variants contribute to ASD risk (45,46). Notably, whole-exome sequencing studies have identified hundreds of genes, with rare and mostly de novo variants conferring a relatively higher risk of ASD (47,48). Their molecular pathways have been largely grouped into those serving synaptic function, chromatic modification, or transcriptional modulation. As a result, genetics-first approaches, whereby specific genotypes represent the target unit for subtyping, have arisen (49–51). Although promising, this approach faces several barriers. These include the different penetrance between common and rare variants, the low prevalence of any single variant/mutation [1%–2% of ASD cases (52), although cumulatively all identified ASD-related gene mutations contribute to ~20%–40% of clinical nonsyndromic ASD], the high genotype-to-phenotype variability, and variable degrees of spatiotemporal convergence across independent risk genes (32). Thus, efforts focused on discovering clinically relevant subtypes based on genetic approaches alone can be challenging.

Recognizing these challenges, researchers have begun to explore the utility of applying data-driven subtyping strategies to complementary biological units of analysis, bringing into focus a range of biophysical indices of brain function [e.g., (53)]. Among them is neuroimaging, a noninvasive tool used to probe brain structural and functional macroscale organization, which in turn is intermediately positioned between behavioral and microscale biological units of analysis (Supplemental Figure S1). Accumulating evidence has suggested the potential utility of neuroimaging in capturing atypical brain organization in ASD (54–58), although with findings varying across studies (59–63). While variation in findings can be attributed to methodological confounds (e.g., sample size, motion artifacts, measurement noise), there is a growing awareness of the potential contribution of neurobiological variation in the neural correlates of ASD. This has motivated the emergence of studies aimed to identify homogeneous ASD subgroups based on neuroimaging features (hereafter referred to as neurosubtyping). This review selectively focuses on this nascent data-driven neurosubtyping ASD literature, summarizing its converging themes, challenges, and potential solutions.

DATA-DRIVEN ASD NEUROSUBTYPING

The current ASD neurosubtyping literature is in its infancy, with a total of 12 studies in humans, 92% of them published since 2018 (64–75). Studies vary in methodology (Figure 1 and Table 1). As such, we stratify this review based on key methodological components (Figure 2), with each section highlighting existing challenges and possible solutions. A summary of converging results follows. Briefly, across studies, findings revealed the presence of 2 to 4 ASD neurosubtypes. Each represented different proportions of subjects, suggesting that traditional case-control group means can miss information about the less represented subtype(s) in a given sample. This has likely contributed to the apparent inconsistencies in traditional group comparisons (10). Most studies associated neurosubtypes with clinical behavioral features that support their likely clinical validity. Finally, as discussed below, the spatially distributed nature of results suggests the relevance of a system neuroscience approach.

Diagnostic Samples

Current Literature. Of the 12 ASD neurosubtyping studies, 7 (59%) neurosubtyped individuals with a confirmed ASD diagnosis separately from neurotypical (NT) individuals and 3 examined ASD and NT data together. Only 2 studies examined transdiagnostic samples. Solely focusing on ASD allows the parsimonious detection of ASD neurosubtypes that may serve as clinical biomarkers. However, because variability also exists within NT individuals, at the very least some standardization of ASD data is recommended to determine the extent to which findings are ASD specific. Three studies normalized the metrics against NT data before neurosubtyping [i.e., structural features (66,69) and *z*-scores of intrinsic functional connectivity (iFC) (75)]. Others (64,71,72) opted to identify neurosubtypes across both ASD and NT datasets, with the assumption that ASD-related pathophysiology may occur in addition to, and interacting with, differing neurosubtypes in the population. Building on these considerations, and consistent with transdiagnostic perspectives in psychiatry (1,76), 2 neurosubtyping studies combined ASD with other psychiatric conditions—schizophrenia spectrum disorders, bipolar disorder, and NT individuals in (71) and attention-deficit/hyperactivity disorder and NT individuals in (72).

Challenges and Possible Solutions. Findings from these early transdiagnostic studies (71,72) suggest that common sources of neurophenotypic variation in ASD extend across diagnostic boundaries (Table 1). For example, Stefanik *et al.* (71) identified 4 neurosubtypes, each including multiple diagnoses at different prevalence and characterized by distinct patterns of anatomical features. Transdiagnostic commonality can result from different scenarios, including additive or pleiotropic mechanisms. These scenarios may be disentangled via widely and deeply characterized large samples to examine continuous psychological and biological dimensions. Such endeavors require significant resources that are, however, taxing for single sites (77). Currently, sample sizes in ASD

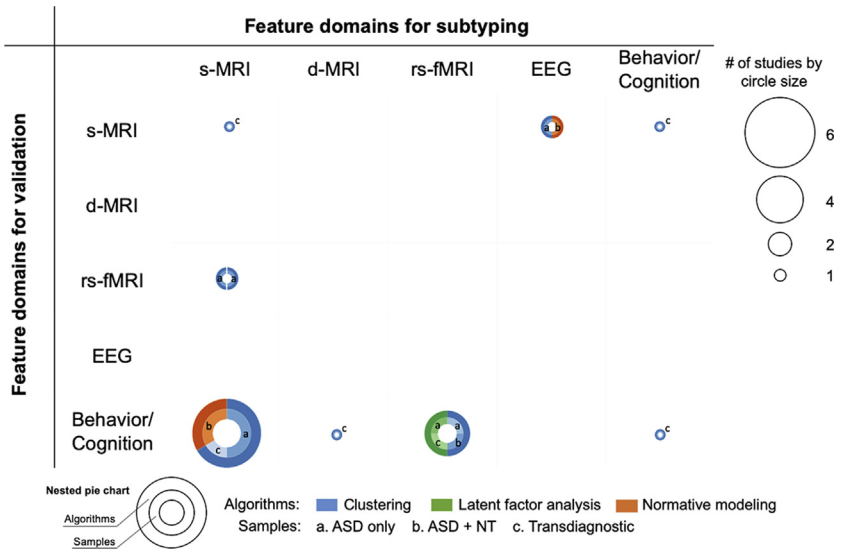

Figure 1. Minimally overlapping neurosubtyping research designs in autism spectrum disorder (ASD). Research designs across current ASD neurosubtyping studies are displayed in matrix form. Columns denote the feature domains used for a given subtyping algorithm. Rows denote the feature domains used for validation. The nested pie charts present further information about variability of subtyping approaches in terms of employed algorithms (outer pie layer; blue for clustering, green for latent factor analysis, and orange for normative modeling) and samples (inner pie layer; letter *a* for individuals with ASD only, letter *b* for individuals with ASD + neurotypical individuals (NT), and letter *c* for transdiagnostic cohorts). The size of the pie chart refers to the number of existing studies at a given feature domain junction (matrix index). d-MRI, diffusion-based magnetic resonance imaging; EEG, electroencephalogram; rs-fMRI, resting-state functional magnetic resonance imaging; s-MRI, structural magnetic resonance imaging.

neurosubtyping studies range from N = 44 to N = 900+ individuals. The largest samples result from retrospective data aggregation and from prospective multicenter collections—the Autism Brain Imaging Data Exchange [ABIDE I and II (78,79)] and the EU-AIMS (European Autism Interventions—A Multicentre Study for Developing New Medications) Longitudinal European Project (80), respectively. Table 2 summarizes neuroimaging datasets (50,81–83) usable for discovery and/or replication [see (84)].

Features

Current Literature. All studies have focused on a single neuroimaging modality, primarily structural or functional magnetic resonance imaging, except 2 electroencephalogram (EEG) studies (65,73). Within imaging modality, the type and number of features varied (e.g., iFC, EEG coherence, cortical thickness, mean diffusivity) (Table 1). High variability of features reflects the challenge of establishing an optimal set of variables and the lack of consensus on what features are most relevant for ASD. Features were predominantly selected based on data availability and/or prior knowledge. In one example, indices of cortical organization were selected based on previous postmortem histological and in vivo neuroimaging findings (Supplemental Figure S2A) (85–87). Only 1 study used a quantitative method for feature selection (65), whereby the sample was clustered among 40 EEG coherence factors selected based on their accuracy in discriminating individuals with ASD from NT individuals (65). This approach can be categorized as a filter method (88,89)—a group of methods selecting features based on their relevance, here to ASD. Although computationally efficient and scalable, most filter methods assess each feature separately, thereby ignoring interfeature dependencies. Although their independence from the subtyping algorithm can be a strength, it may also lead to selecting suboptimal features (88). To address this limitation, other quantitative approaches select features based on their usefulness to the specific question of interest; these are referred to as wrapper (e.g., sequential search) or embedded (e.g., random forest) approaches [see (88,89) for reviews]. These methods are yet to be applied to ASD neurosubtyping, although promising proofs of concept are discussed below (53,90).

One critical consideration for feature selection is data reduction. Although deeply characterized samples with a high number of features to explore across multiple modalities are theoretically desired, data with more feature dimensions than subjects create a statistically ill-posed problem—the fact that an infinite number of solutions can fit data equally well. Many algorithms rely on distance metrics quantifying how far the data points are from each other. As the dimensionality gets larger, the interpoint distance becomes less discernible (91,92). To date, only two ASD neurosubtyping studies have quantitatively addressed this challenge—one via principal component analysis (PCA) (65) and one via nonnegative matrix factorization (66).

A related question is how to optimally combine different data modalities. Current ASD studies have focused on single neuroimaging modalities, and none has combined neuroimaging with behavioral or other biological units of analysis. In one exception (71), similarity network fusion (93), a variant of multiview multidimensional scaling (94), was applied transdiagnostically to combine demographic, brain imaging, and behavioral data (Supplemental Figure S2B). After detecting intersubject distances across all pairs of data domains, this approach jointly embedded those distances in a lower dimensional subspace. This process resulted in 4 transdiagnostic community structures showing distinct neurocognitive profiles. Although this approach is helpful, caveats of similarity network fusion include its limited scalability to highly dimensional data (95) and sensitivity to noise given that it assigns equal weights across features (96).

Table 1. ASD Neurosubtyping Studies in Humans

	Sample[a]				Approach			Results	
Study	Data Source	*N*	Age Range,[b] Years	Male Individuals, *n* (%)	Subtyping	Feature(s)	Validation[c]	No. of Neurosubtypes	Highlights
(64)	ABIDE I	145 ASD 121 NT	7–39	All (100)	Clustering (*k*-means)	R-fMRI; iFC 160 cortical ROIs[d]	SRS, IQ, ADOS	2	NS1 (59% ASD, 45% NT), NS2 (41% ASD, 55% NT). NS2 had ↓ iFC between networks and ↑ iFC within networks vs. NS1; NSs did not differ in behavior, demographics, IQ, scan performance, or parameters. PLS brain-behavior analyses showed iFC correlations with a combination of symptom scores unique to each NS.
(66)	ABIDE I, II	356 ASD	5–35	All (100)		sMRI; whole-brain VBM	ADOS, iFC	3	NS1 (18%): ↓ prefrontal GMV; NS2 (53%): ↑ temporal and ↓ prefrontal and occipital GMV; NS3 (29%): ↑ GMV temporal vs. NT. ADOS scores were different between NSs. iFC in frontoparietal network differed between NS1 and NS3. NSs improved supervised classification.
(67)	Lab specific	57 ASD	9–18	47 (82)		R-fMRI; occipital to frontal pole cortex iFC	IQ, ADI-R, ADOS, comorbidities, medication use, age, sex	2	Post hoc clustering of iFC in ROIs identified by group mean. NSs had opposite iFC patterns and did not differ in clinical and demographic metrics.
(65)	Lab specific	430 ASD	2–12	361 (84)	Clustering (hierarchical and *k*-means)	EEG; 40 coherence factors	NA	2	NS1 = 39%; NS2 = 61%. Of the 40 EEG features, 19 discriminated NS1 (↑ coherence in 7 of 40 factors; ↓ coherence in 12 of 40 factors) vs. NS2 (opposite patterns with NS1) and both NSs vs. NT. NSs did not differ for age or sex.
(69)	ABIDE I	107 ASD	7–50	All (100)		sMRI; CT, SA, IC, GD	ADOS, iFC	3	NS1: ↑ CT, SA, TB; NS2: ↓ CT, GD; NS3: ↑ GD. Distinct ADOS and iFC load across NSs. NSs improved supervised ADOS prediction above NS blind approach.
(68)	Lab specific	64 ASD	9 ± 6	52 (81)		sMRI; 2D-slice CT, volumes in 6 ROIs[d]	EEG profile, epilepsy, medical history, CARS	4	NS1 (28%; ↑ CC); NS2 (52%; ↑ amygdala and hippocampus); NS3 (14%; ↑ NC and ↓ hippocampus); NS4 (6%; ↓ CC, amygdala, NC). NSs did not differ by age, IQ, or severity but differed in ROI atypicalities, pregnancy order, and psychomotor early delay (NS4 > NS2)

Table 1. Continued

Study	Sample[a]				Approach			Results	
	Data Source	*N*	Age Range,[b] Years	Male Individuals, *n* (%)	Subtyping	Feature(s)	Validation[c]	No. of Neurosubtypes	Highlights
(71)	Lab specific	38 ASD 34 BD 51 SSD 51 NT	16–35	26 (76) ASD 19 (56) BD 32 (63) SSD 24 (47) NT	Spectral clustering (SNF)	sMRI; demographics, cognition CT, subcortical volume, DTI-FA, MD	Cognitive and emotion test scores, CT-based global efficiency	4	NS1 (80% ASD followed by SSD and BD;↑ SCV, ↓ FA, ↑ CT); NS2 (64% SSD and minor distribution of other diagnoses; ↓ SCV, ↓ FA, ↑ CT); NS3 (46% NT and ∼20% each of the others; ↓ SCV, ↓ FA, ↓ CT); NS4 (similar proportions across diagnoses; ↑ SCV, ↑ FA).
(72)	ABIDE I and ADHD-200	369 ASD 284 ADHD 652 NT[e]	7–21	All (100)	LFA[f]	R-fMRI; iFC of 21 ROIs in DN, SN, DAN[g]	Diagnostic labels, symptom questionnaires (unspecified)	3	NS1 = ↑ DN–DAN, medium DN–SN, ↓ intra-DN and intra-DAN iFC, and positive association with ADHD; NS2 = ↓ DN–DAN and DN–SN, associated with ASD diagnosis, language, and IQ; NS3: ↓ DN–DAN with no behavioral associations.
(75)	ABIDE II and GENDAAR	306 ASD[h]	15 ± 6/ group	236 (77)	LFA[f]	R-fMRI; 418 cortical + subcortical iFC[i]	Multiple clinical and demographics	3	NSs had dissociable whole-brain hypo/ hyper iFC and shared atypical iFC in DN. Individuals expressed multiple NSs at different degrees. NS1 associated with ↓ FC (DAN, SM, SN, VN) and ↑ FC (DN, CO) in ASD and with symptom severity, NS2 with opposite patterns of FC (compared with NS1) and with comorbid symptoms. NS3 (complex mixture of ↑ and ↓ FC) preferentially expressed in older male individuals.
(74)	EU-AIMS LEAP	321 ASD 206 NT	6–31	232 (72) ASD 127 (62) NT	NM[j]	sMRI; age- and sex-related CT, SA	ADOS, ADI-R	2	28% of the ASD cohort had regional age-related CT deviations from the normative curve vs. 19% in NT. ↑ deviation in ASD vs. NT. CT deviation associated with RRB severity in female individuals. Common case-control comparisons yielded minimal regional differences in contrast to those with the outlier subgroup.

Table 1. Continued

Study	Sample[a]				Approach			Results	
	Data Source	*N*	Age Range,[b] Years	Male Individuals, *n* (%)	Subtyping	Feature(s)	Validation[c]	No. of Neurosubtypes	Highlights
(70)	ABIDE I, II	942 ASD[k]	6–20	754 (80)	NM[l]	sMRI; age-related CT	NA	2	7% to 10% of ASD cohort had age-related CT deviance (>2 w-score)[m]. Case-control group comparison revealed small effects and spatially limited differences.
(73)	Lab specific	44 ASD 44 NT	10 ± 4/ group	33 (75)/group		EEG; age-related alpha wave	sMRI volume	2	↑ magnitude of alpha waves deviation in ASD vs. NT. Correlation between regional volume and alpha frequency did not survive statistical correction. Case-control did not yield any group differences.

ABIDE, Autism Brain Imaging Data Exchange; ADHD, attention-deficit/hyperactivity disorder; ADI-R, Autism Diagnostic Interview–Revised; ADOS, Autism Diagnostic Observation Schedule; ASD, autism spectrum disorder; BD, bipolar disorder; CARS, Childhood Autism Rating Scale; CC, corpus callosum; CO, control; CT, cortical thickness; DAN, dorsal attention network; DN, default mode network; DTI, diffusion tensor imaging; EEG, electroencephalogram; EU-AIMS LEAP, European Autism Interventions Longitudinal European Autism Project; FA, fractional anisotropy; FC, functional connectivity; GD, geodesic distance; GENDAAR, Gender Exploration of Neurogenetics and Development to Advance Autism Research; GMV, gray matter volume; IC, intensity contrast; iFC, intrinsic functional connectivity; LFA, latent factor analysis; MD, mean diffusivity; NA, not applicable; NC, nucleus caudatus; NM, normative modeling; NS, neurosubtype; NT, neurotypical; PLS, partial least squares; R-fMRI, resting-state functional magnetic resonance imaging; ROI, region of interest; RRB, restricted repetitive behavior; SA, surface area; SCV, subcortical volume; SM, somatomotor; sMRI, structural MRI; SN, salience network; SNF, similarity network fusion; SRS, Social Responsiveness Scale; SSD, schizophrenia spectrum disorders; TB, tissue blurring; VBM, voxel-based morphometry; VN, visual network.

[a]Only the sample being neurosubtyped is reported.

[b]Mean age and standard deviation are used when age range was not reported.

[c]Reporting validation based on domains distinct from the features originally used to identify NSs.

[d]ROIs were based on the Dosenbach atlas (171).

[e]303 NT from ADHD-200, 349 NT from ABIDE I.

[f]LFA used a Bayesian model based on latent Dirichlet allocation.

[g]ROIs included nodes of the DN, SN, and DAN selected based on meta-analyses.

[h]NT sample (n = 348) from ABIDE II + GENDAAR was used to generate iFC z-scores in ASD.

[i]ROIs included the 400 cortical parcellations per Schaefer *et al*. (172) and 18 subcortical FreeSurfer parcellations.

[j]NM was based on Gaussian process regression.

[k]n = 870 NT data were used to build the normative model.

[l]NM was based on local polynomial regression.

[m]>2 w-score is a statistical norm (analogous to a z-score) based on a mean and standard deviation from the TD group that is age- and sex-matched.

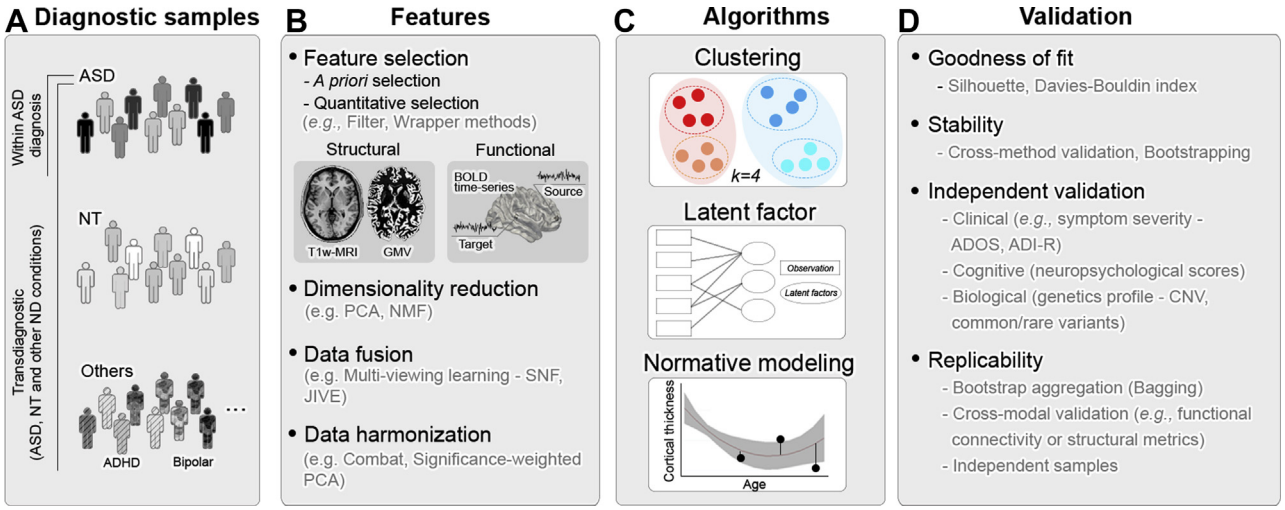

Figure 2. Key methodological steps in neurosubtyping. **(A)** Selection strategies for diagnostic samples may focus on neurosubtyping within the autism spectrum disorder (ASD) or extend it to neurotypical individuals (NT) and/or those with other psychiatric diagnoses. Including NT can account for variance in the neurotypical population, while including other psychiatric conditions allows examining the extent to which neurosubtypes are specific to a given diagnosis. **(B)** Several steps exist for feature selection and processing. Feature selection (whether in the brain structural domain or functional domain) may be based on a priori knowledge or on quantitative approaches (88). Critical steps for feature processing include dimensionality reduction, data harmonization, and multimodal feature combination. **(C)** The current ASD neurosubtyping literature comprises clustering-based analysis, latent factor analysis, and normative modeling. See Table 1 for details. **(D)** Primary strategies for validating neurosubtyping results. See Table 1 for details. ADHD, attention-deficit/hyperactivity disorder; ADI-R, Autism Diagnostic Interview–Revised; ADOS, Autism Diagnostic Observation Schedule; BOLD, blood oxygen level–dependent; CNV, copy number variation; EEG, electroencephalogram; GMV, gray matter volume; iFC, intrinsic functional connectivity; JIVE, joint and individual variation explained; PCA, principal component analysis; ND, neurodevelopmental disorders; NMF, nonnegative matrix factorization; SNF, similarity network fusion; T1w-MRI, T1-weighted magnetic resonance imaging; VBM, voxel-based morphometry.

Challenges and Possible Solutions. To enable quantitative selection of those features most tied to the outcome of interest, two hybrid approaches are promising for ASD neurosubtyping. One is functional random forest (FRF) (53), and the other is surrogate variable analysis (SVA) (90). FRF combines random forest (supervised) with community detection (unsupervised). Briefly, FRF generates an ensemble of decision trees where each leaf node contains individuals who are highly similar with respect to a variable of interest. This information constructs a subjectwise proximity matrix, which in turn is fed into community detection. FRF has recently identified subtypes based on cognitive performance in children with ASD and NT children (90). In contrast, SVA has not yet been applied to ASD or subtyping. Originally developed to remove unknown sources for batch effects in genomic data (90), SVA identifies and decomposes confounding sources combining linear regression (supervised) and PCA (unsupervised). A common caveat of both methods is that one needs to select the appropriate relevant question for subtypes to be informative (97). In ASD, examples of relevant questions pertain to whether neuroconnectome subtypes reflect distinct genetic profiles, distinct treatment responses, or different comorbidity risks. Future applications across multiple units of analysis are required to confirm FRF and SVA usefulness in ASD.

In relation to multiview learning methods (98) that address the scalability limitation of similarity network fusion discussed above, notable examples include generalized canonical correlation analysis and joint and individual variation explained (JIVE) (99). Generalized canonical correlation analysis is a generalization of PCA and linear regression to more views of the data. It jointly learns a low-dimensional representation of each view (i.e., similarity matrix) that maximizes correlation between all pairs of views. An advantage is that it generates differential weights for each feature. JIVE (99) identifies variance components spanning multiple data types and their unique variation. Thus, it provides a comprehensive description of multisource heterogeneity within and across data types. Both techniques have yet to be used in neurosubtyping, but recent neuroimaging studies support their utility in predicting the brain's relationship with age and behavior (100,101) as well as in extracting a low-dimensional representation of depression-related connectivity (102).

For any feature selection, additional problems should be addressed. One is data leakage (103), which occurs when feature selection and subtyping are dependent, thereby artificially amplifying the quality of subtyping solution (i.e., overfitting the model). Any algorithm based on strict cross-validation, such as embedded methods, bypasses this risk by splitting the data into training and test cases. Another set of problems is related to employing features that can be driven by artifacts (e.g., head motion) and/or exhibit inadequate reliabilities [e.g., across scans, sessions, and/or scanners (104)]. These raise concerns about the "garbage in, garbage out" concept (i.e., low-quality features yield invalid and/or unreproducible results). Head displacements as low as 0.2 mm are the most critical artifact for high–spatial resolution neuroimaging studies because they can systematically affect findings from functional, structural, and diffusion neuroimaging (105–113). In response, many labs have developed robust approaches that limit motion in data collection (106,114,115)

Table 2. Selected Imaging Datasets Suitable for ASD Neurosubtyping Discovery and Replication

Source	Descriptor	Design	Sample *N*	Age Range, Years	Imaging Modality	Phenotypic and Other Data	Open	Ongoing	Neurosubtyping Use
ABIDE I[a]	Previously collected neuroimaging and phenotypic data across 17 sites using independent protocols. Data released in August 2012 (78)	CS	539 ASD, 573 NT	7–64	sMRI, R-fMRI	Demographics, ADOS, IQ, partially enhanced (e.g., ADI-R comorbidity measures) in some sites	Yes	No	Multimodal neurosubtyping within ASD diagnosis or across both ASD and NT. Existing but limited non-neuroimaging data available for validation or data fusion. Open sharing facilitates independent replications of findings in more deeply phenotype datasets.
ABIDE II[b]	Previously collected imaging and phenotypic data across 19 sites using independent protocols. Data released in June 2016 (78)	Largely CS (longitudinal in 2 sites)	521 ASD, 593 NT	5–64	sMRI, R-fMRI, (DTI for selected sites)	Partially enhanced phenotypic information relative to ABIDE I	Yes	No	
NDAR[c]	Retrospective and ongoing collections from multiple independent studies/protocols (82)	Varying by project, include CS and longitudinal	Varying by project, ASD, NT, and other diagnoses	Varying by study	Multimodal imaging, varying by study	Psychiatric, other phenotypes; genetics and other metrics varying by study	Yes	Mixed	Multimodal neurosubtyping within ASD diagnosis or across both ASD and NT and/or multiple diagnoses. Multidimensional phenotypic data available for validation along with other objective markers, but availability is highly variable across studies. Open sharing facilitates independent replications of findings, may need independent, more deeply phenotypic datasets.

Table 2. Continued

Source	Descriptor	Design	Sample *N*	Age Range, Years	Imaging Modality	Phenotypic and Other Data	Open	Ongoing	Neurosubtyping Use
EU-AIMS LEAP[d]	Large-scale prospective multicenter (7 sites) data collection (2014–2016) (80)	Longitudinal (61% of baseline data completed follow-up assessments)	437 ASD, ~300 NT and mild ID, 36 ASD twins, 36 NT twins	6–30	sMRI, FLAIR, DTI R-fMRI, T-fMRI, EEG	Clinical symptom, comorbidities, quality of life, neurocognitive, biochemical, prenatal environmental risk factors and genomics	No	Yes	Multimodal neurosubtyping within ASD diagnosis or across both ASD and NT and/or ID. ASD-specific multidimensional phenotypic data available for validation or data fusion along with other objective markers (e.g., eye tracking). Longitudinal data in >61% of cases, allow to use behavioral outcomes for validation and brain trajectories for subtyping.
HBN[e]	Community self-referred multiomic dataset, aimed at 10,000 individuals from the New York City metropolitan area (173)	CS	Transdiagnostic (16% ASD), NT; to date, ~3000 data collected, ~2000 openly released	5–21	sMRI, R-fMRI, naturalistic viewing fMRI, T-fMRI, EEG	Psychiatric, behavioral, cognitive, and lifestyle phenotypes; eye-tracking, voice and video-recordings, genetics, and actigraphy	Yes	Yes	Multimodal neurosubtyping within ASD diagnosis or across both ASD and NT and/or multiple diagnoses. Extended multidimensional phenotypic datasets available for validation or data fusion along with other objective markers (e.g., eye tracking). Open sharing facilitates independent data replications also using split samples.
POND[f]	A clinical multidiscipline study of the neurobiology of multiple neurodevelopment disorders (174)	CS; longitudinal treatment in subsamples	Transdiagnostic (ASD, OCD, ADHD, ID, unknown genetic syndromes)	School age	sMRI, R-fMRI, T-fMRI, MEG	Psychiatric, behavioral, cognitive	No	Yes	Multimodal neurosubtyping within ASD diagnosis or across diagnoses. Multidimensional phenotypic dataset available for validation.

Table 2. Continued

Source	Descriptor	Design	Sample *N*	Age Range, Years	Imaging Modality	Phenotypic and Other Data	Open	Ongoing	Neurosubtyping Use
NKI-RS[g]	Community-ascertained lifespan sample (83)	CS; longitudinal	1000 largely NT	6–65	sMRI, R-fMRI	Physiological, psychological, genetic data	Yes	Mixed	Multimodal neurosubtyping of NT variance across multiple psychological domains including ASD-associated traits. The restricted number of individuals with ASD limits the assessment of ASD specificity. Longitudinal outcome available for validation.
Generation R, Generation R Next[h]	Population-based prospective multiethnic cohort from the Rotterdam (Netherlands) metropolitan area Eligibility: pregnant women with an expected delivery date between 2002 and 2006; long-term follow-up design (81)	Longitudinal	Largely NT; a small proportion of children have ASD confirmed at 6 to 9 years of age based on review of medical records	Fetal-16[i]	Ultrasound (prebirth); sMRI, DTI, R-fMRI collected after 6 years of age	Health outcome, environmental, endocrine, genomic, lifestyle, nutritional, sociodemographic determinants	No	Yes	Multimodal neurosubtyping of NT variance across multiple psychological domains including ASD-associated traits. The restricted number of individuals with ASD may limit the assessment of ASD specificity. Extensive longitudinal data offer outcome for neurosubtyping validation.

ABIDE, Autism Brain Imaging Data Exchange; ADHD, attention-deficit/hyperactivity disorder; ADI-R, Autism Diagnostic Interview–Revised; ADOS, Autism Diagnostic Observation Schedule; ASD, autism spectrum disorder; CS, cross-sectional; DTI, diffusion tensor imaging; EEG, electroencephalogram; EU-AIMS LEAP, European Autism Interventions Longitudinal European Autism Project; FLAIR, fluid-attenuated inversion recovery image; HBN, Healthy Brain Network; ID, intellectual disability; MEG, magnetoencephalography; NDAR, National Database for Autism Research; NKI-RS, Enhanced Nathan Kline Institute–Rockland Sample; NT, neurotypical; OCD, obsessive-compulsive disorder; POND, Province of Ontario Neurodevelopmental Disorders; R-fMRI, resting-state functional magnetic resonance imaging; sMRI, structural magnetic resonance imaging; T-fMRI, task-based functional magnetic resonance imaging.

[a] http://fcon_1000.projects.nitrc.org/indi/abide/abide_I.html.
[b] http://fcon_1000.projects.nitrc.org/indi/abide/abide_II.html.
[c] https://nda.nih.gov/about.html.
[d] https://www.eu-aims.eu.
[e] http://fcon_1000.projects.nitrc.org/indi/cmi_healthy_brain_network/.
[f] https://pond-network.ca.
[g] http://fcon_1000.projects.nitrc.org/indi/enhanced/.
[h] https://generationr.nl/.
[i] Sample followed longitudinally before broth and up to 16 years of age.

as well as detect and analytically mitigate their remaining effects [e.g., scrubbing, independent component analysis–automatic removal of motion artifacts (106,116)].

Furthermore, as in other neuroscientific disciplines, the neuroimaging field has at times forged ahead with features that were not properly assessed or optimized for reliability. This is problematic because with the exception of core structural measures (e.g., volume, cortical thickness), most neuroimaging metrics have modest to moderate reliability (104). This increases the sample size needed to identify meaningful group-level findings and severely limits opportunities for individual-level analysis (117). Current ASD neurosubtyping studies have examined some features with moderate to high test–retest reliability [e.g., cortical thickness: intraclass correlation coefficient = 0.6–0.8 (118,119)]. For others, reliability remains unknown (e.g., gray/white matter intensity contrast). Efforts to optimize data acquisitions and analyses to improve reliability are underway and should be considered in ASD neurosubtyping. For example, in functional connectomics, increasing data acquisitions from ~5 to 30 minutes per subject has been shown to dramatically increase reliability even if combining data across differing functional magnetic resonance imaging scan types [e.g., task, rest (120,121)].

Algorithms

Current Literature. Data-driven neurosubtyping in ASD has comprised three methods. Two of these methods include fully unsupervised approaches, such as clustering (n = 7 studies) and latent factor analysis (n = 2), and a third is normative modeling (n = 3).

Clustering partitions the data into a number of clusters such that the samples in each cluster are more similar to one another than to those in other clusters. Reflecting the widespread popularity and relative computational ease of k-means clustering—a centroid-based approach that iteratively assigns individual data points based on their closeness to the centroid of each cluster—3 ASD neurosubtyping studies (64,66,67) have used this type of clustering. Others (65,68,69) have used variants of hierarchical clustering that generate cluster trees using heuristic data splitting or merging (Supplemental Figure S2A). Unlike k-means, hierarchical clustering does not require specifying cluster numbers beforehand. With hierarchical clustering, each branch of the tree tracks a sequence of progressive clustering processes, fully capturing the nested hierarchical data structures. This facilitates interpretation of neurosubtyping across multiple scales because different subgroup resolutions can guide clinical decisions. On the other hand, k-means needs to fit only a few boundaries to data, while hierarchical clustering learns many more boundaries, potentially leading to suboptimal findings.

Regarding latent factor analysis, 2 recent studies (72,75) have used this approach by borrowing topic modeling techniques from natural language processing (Supplemental Figure S2C). These models assume that each document (individual brain) is a collection of words (neuroimaging feature) associated with a subset of k latent topics (k brain factors). They automatically summarize high-dimensional individual data into a combination of parsimonious latent structures. In this way, they allow each individual to express multiple latent brain factors to differing degrees, which potentially reflects heterogeneous underlying pathologies. Thus, the advantage of this approach is that it quantitatively describes both categorical and dimensional aspects of ASD heterogeneity.

Regarding normative modeling, by estimating the quantiles of variation for a brain metric in a given population, normative modeling statistically infers each individual deviation from the normative pattern (Supplemental Figure S2D) (122). The identification of the extreme cases in the distribution is used to find a subgroup of individuals who are clearly more separable from all others and possibly more homogeneous. To date, 3 ASD neuroimaging studies (70,73,74) have adopted this approach. Focusing on either cortical thickness (70,74) or alpha wave EEG profile deviation from normative curves derived from cross-sectional age data (73), the results of these studies showed that variability between individuals with ASD and NT individuals is largely overlapping and only a subset of individuals showed a high degree of brain atypicalities. This method does not assume that there are multiple subgroups. Conversely, it highlights a deviation score for each individual rather than a group mean—a concept in line with highly desired precision medicine (4). A caveat of this approach is that estimating high-dimensional distributions requires an exponentially increasing number of data points as the number of dimensions increases. Because this cannot be practically done in the native dimensionality of neuroimaging data, a smaller number of features must be selected relative to other algorithms that do not need to estimate the full distribution but only need to estimate partitions of the space. Therefore, choosing the right features is more important for normative modeling.

Validation

Current Literature. Regardless of the specific feature or subtyping algorithm employed, within a given study, a range of assumptions are made and parameters are estimated. As such, there is no single one-size-fits-all subtype solution. Different approaches have been developed to directly estimate optimal solutions within the same dataset [see (123) for a review]. Some approaches have been used in current ASD neurosubtyping [e.g., silhouette (66), Davies–Bouldin index (69)]. Others indirectly validate the subtype solutions by assessing their stability either across independent subtyping algorithms or via bootstrapping. So far, only 2 ASD neurosubtyping studies have assessed the convergence of findings from different algorithms (65,75) and 4 have used bootstrapping (64,69,70,75); all reported within-study stability.

Different solutions can be equally meaningful, with each solution capturing a distinct aspect of the data (124). Thus, a common within-sample strategy for validation is to demonstrate that identified neurosubtypes explain variation in measures other than those indexing the subtyping features. In the current ASD neurosubtyping literature, such a validation method has been primarily based on clinical and demographic metrics. Although preliminary, initial results suggest that the identification of neurosubtypes may be promising to explain brain–behavior relationships in regard to symptom severity (69) or diagnostic labels (66) better than group average results. Four studies have used independent neuroimaging modalities (66,69,71,73), whereby neurosubtypes identified via regional

structural features were validated with brain connectivity metrics. These studies point toward neurosubtype-specific associations with regional and large-scale brain organization.

Challenges and Possible Solutions. An existing gap in clinically relevant validation of ASD neurosubtypes is to assess their relation to clinical outcomes. The utility of this approach has been shown in Alzheimer's disease and depression, whereby neurosubtype-specific outcomes have been revealed based on naturalistic prospective and treatment data, respectively (102,125). Longitudinal datasets also allow assessing the stability of a subtype solution over time; developmentally stable subtypes are likely to represent heterogeneity of disease traits. In contrast, unstable subtypes may reflect variation in developmental stages and/or disease progression. Unfortunately, in ASD no longitudinal data have been used in data-driven neurosubtyping. This likely reflects their limited availability, which should be a mandatory future focus (126). Meanwhile, preliminary investigation can be carried out cross-sectionally.

In addition, mostly owing to compliance requirements for neuroimaging data collection, the field has largely failed to include less cognitively capable individuals, thereby limiting the generalizability of findings. Future studies must invest in protocols that allow data collection across all abilities, including natural sleep scans (127,128), passive viewing (129,130), scan behavioral preparation (131), and technical advancements addressing head motion prospectively (114,115,132). Another challenge germane to ASD is its disproportional representation of male individuals (133). While sex-related differences may contribute to ASD biological heterogeneity (134,135), the higher prevalence of male individuals with ASD and the tendency of single-site studies to exclude, or minimally represent, female subjects may limit generalizability of neurosubtypes. Fortunately, prospective studies explicitly overrecruiting female subjects are ongoing (136), and large-scale data repositories are amassing larger amounts of female data (79).

Like any clinically useful biomarkers, high precision (low false-positive rate), recall (low false-negative rate), and (most important) replicability (the degree to which identical findings are obtained in a distinct sample with similar methods) are key requirements (104). Unfortunately, many fields, including neuroscience and psychology, are currently experiencing a replicability crisis (137,138). Two factors contribute to this: sample size and measurement reliability (117). Statistical strategies to increase replicability exist. A representative one is bagging (i.e., bootstrap aggregating), a technique that aggregates randomly selected subsamples to reduce the variability of the measurement through averaging (139). A recent study applying bagging to brain parcellation demonstrated substantially improved reproducibility and test–retest reliability (140). A similar idea can be applicable to neurosubtyping by running the clustering algorithm on thousands of bootstrap samples. Averaging across bootstrap-derived clustering yields a consensus matrix (141,142) representing how consistently a pair of individuals was subgrouped together. This comes with high cost in computation and possibly in interpretability. One more general caveat is that bagging cannot overcome limitations of the sample population; for example, if it is a biased sample, naive bagging will not reduce the bias. More sophisticated stratified resampling techniques can mitigate these biases.

Replicability is also affected by high variance coming with large multisite datasets. This so-called batch effect can be attenuated by harmonization processes, namely the PCA-based data reconstruction method (143) or ComBat [combining batches (144)]. ComBat, originally developed in the genomic field, is based on a linear model involving the site as a main (additive and multiplicative) statistical term. The model reduces the site effects based on statistical correction and reconstructs a minimized site effect dataset [see (145) for discussion on residual site effects]. Compared with conventional regression-based corrections, ComBat is more robust to outliers in small samples because it uses an empirical Bayes approach to estimate site/time collection effects. Significance-weighted PCA is another mitigation approach that involves performing PCA over the whole images and then computing the statistical significance of each component in relation to a site. An initial application successfully obtained site-variable weighted subnetworks in an ASD nonsubtyping study (143). To facilitate replication, we draw attention to existing neuroimaging ASD data repositories in Table 2.

CONVERGING FINDINGS, FUTURE AVENUES

Although the ASD neurosubtyping literature to date is preliminary, some convergence in findings begins to emerge. First, most of the ASD neurosubtypes so far identified a combination of ASD-related increases and decreases in any neuroimaging feature examined. Second, regardless of the features examined, no ASD neurosubtype has been characterized by a spatially isolated (focal) pattern; instead, atypicalities are spatially distributed. In this regard, a set of regions within the default mode and frontoparietal networks are consistently involved across neurosubtypes (Supplemental Figure S3). Given that higher-order social and cognitive processes are affected in ASD, we suggest that the common involvement of these networks may underlie shared impairment in core ASD symptoms. The subtype-specific pattern of atypicality within these networks (i.e., increases and/or decreases in a given feature) may differentially affect symptom severity and presentation. On the other hand, the degree to which atypicalities involve other networks more sporadically observed across neurosubtypes (e.g., visual and dorsal attention networks) may be associated with interindividual differences in comorbid symptoms. Testing this hypothesis requires combining neuroimaging with phenotypic characterization of both ASD core and comorbid symptoms, which is uncommon in ASD neuroimaging, with few exceptions (79,146–149).

The distributed nature of neural atypicalities in ASD neurosubtypes points toward mechanisms affecting large-scale brain organization. Thus, measures of brain connectivity (e.g., iFC, structural covariance, EEG coherence) may more directly guide toward the biology underlying ASD heterogeneity. A range of models of atypical connectivity in ASD exists, including decreased long range versus increased short range

(150), imbalanced intramodular versus intermodular (151), atypical shifts toward cortical–subcortical (78), and idiosyncratic connectivity (152). More recently, atypical cortical connectome hierarchy has been proposed as a summary statistic to recapitulate most of the patterns above (153). Testing these models in the context of neurosubtyping may provide greater insight into their biological relevance in ASD heterogeneity.

Given what has been learned from advances in genetics, it should be acknowledged that there is no guarantee that neurosubtypes will directly map onto behavioral heterogeneity. Indeed, recent approaches suggest that distinct single-gene mutations (or subsets of them) may lead to the same behavioral outcomes by converging at one or multiple levels of analysis going from proteins to macroscale circuits. Each level in turn may have different spatiotemporal vulnerabilities, thereby further adding complexity (32,154,155). Similarly, it is possible that dysfunctional brain circuits specific to a given neurosubtype may arise from different cellular pathways that in turn may be anchored on distinct genes, each with differential spatial and developmental expressions. Many of the solutions discussed above, including longitudinal designs in multidimensional transdiagnostic approaches paralleled with rigorous analytical methods for data reduction, validity, and reproducibility, will likely facilitate the efforts in decomposing ASD heterogeneity.

A remaining challenge is to bridge macroscale brain phenomenology to microscale underlying mechanisms (i.e., cellular, molecular, and genetic). Although the cognitive and sociocommunicative domains affected by ASD are human specific and cannot be comprehensively recapitulated by any animal model, research using model organisms amenable to genetic manipulation offers one means to bridge different scales. Currently, a macaque model of one genetic variant observed in humans with ASD has been reported (156), and a larger number of models are available in the laboratory mouse (157–159). The application of translationally relevant neuroimaging across these models may reveal neurosubtypes. For example, morphoanatomical brain mapping across 26 mouse ASD genetic models (160) has identified three neurosubtypes affecting distinct sets of brain regions with shared patterns across seemingly unrelated ASD mutations. The recent implementation of resting-state functional magnetic resonance imaging in mice (158,161) also offers the possibility to expand this approach to functional networks. Indeed, a previous study has demonstrated that 16p11.2 microdeletion comparably impairs iFC in both humans and mice (162).

Finally, while data-driven subtyping using biophysical methods other than neuroimaging in ASD to date is limited in numbers, future neurosubtyping approaches would benefit from leveraging findings from the larger literature of candidate biomarkers. For example, eye-tracking studies in toddlers have suggested that atypical preferential fixation to geometric stimuli characterizes a subgroup of toddlers with ASD (163–166). This marker could be used in the context of multimodal clustering approaches (e.g., neuroimaging with eye tracking) and/or for independent validations of neurosubtypes. Similarly, neurosubtyping can be enriched with results from ongoing large-scale efforts (167) assessing the feasibility and robustness of markers such as delayed latency in N170 during inverted face processing (167,168), behavioral video tracking during parent–child interaction (169), and eye fixation and/or pupillometry during social dynamic sceneries (170).

CONCLUSIONS

Advances in neuroimaging and computational science, with evidence that ASD heterogeneity uniquely affects multiple domains and scales, have motivated neurosubtyping. Results from initial efforts illustrate its feasibility and potential utility. They also underscore limitations, including the need for larger datasets, wider and deeper phenotyping, and advanced analytical models that validly capture, with high replicability, the hybrid—categorical and dimensional—nature of ASD heterogeneity.

ACKNOWLEDGMENTS AND DISCLOSURES

This work was supported by grants from the National Institute of Mental Health (Grant Nos. R01MH105506 and R01MH115363 [to ADM] and Grant No. 1R21MH116473-01A1 to [AG]), Canadian Institutes of Health Research (Postdoctoral Fellowship No. MFE-158228 [to S-JH]), National Science Foundation (Grant No. EEC-1707298 [to JTV]), Singapore National Research Foundation (Fellowship, Class of 2017 [to BTTY]), European Research Council (Grant No. 802371 [to AG]), Simons Foundation (Grant No. SFARI 400101 [to AG]), Brain and Behavior Foundation (NARSAD Independent Investigator Grant No. 25861 [to AG]), SickKids Foundation (Grant No. NI17-039 [to BCB]), National Sciences and Engineering Research Council of Canada (Discovery Grant No. 1304413 [to BCB]), Canadian Institutes of Health Research (Grant No. FDN-154298 [to BCB]), Azrieli Center for Autism Research, and Canada Research Chairs program (to BB).

We thank Irene Droney for copyediting the latest version of the manuscript.

ADM receives royalties from the publication of the Italian version of the Social Responsiveness Scale—Child Version by Organization Speciali. All other authors report no biomedical financial interests or potential conflicts of interest.

ARTICLE INFORMATION

From the Center for the Developing Brain (S-JH, MPM) and Autism Center (ADM), Child Mind Institute, New York, and Center for Biomedical Imaging and Neuromodulation (MPM), Nathan Kline Institute, Orangeburg, New York; Department of Biomedical Engineering Institute for Computational Medicine (JTV), Kavli Neuroscience Discovery Institute, Johns Hopkins University, Baltimore, Maryland; Martinos Center for Biomedical Imaging (BTTY), Massachusetts General Hospital, Boston, Massachusetts; Functional Neuroimaging Laboratory (AG), Center for Neuroscience and Cognitive Systems, Istituto Italiano di Tecnologia, Rovereto, Italy; McConnell Brain Imaging Centre (BCB), Montreal Neurological Institute and Hospital, McGill University, Montreal, Quebec, Canada; Department of Electrical and Computer Engineering (BTTY), Center for Sleep and Cognition, Clinical Imaging Research Centre, N.1 Institute for Health, and NUS Graduate School for Integrative Sciences and Engineering (BTTY), National University of Singapore, and Centre for Cognitive Neuroscience (BTTY), Duke–NUS Medical School, Singapore.

Address correspondence to Adriana Di Martino, M.D., Autism Center, Child Mind Institute, 101 East 56th Street, New York, NY 10022; E-mail: Adriana.DiMartino@childmind.org.

Received Jul 23, 2019; revised Mar 25, 2020; accepted Mar 28, 2020.

Supplementary material cited in this article is available online at https://doi.org/10.1016/j.biopsych.2020.03.022.

REFERENCES

1. Insel TR (2014): The NIMH Research Domain Criteria (RDoC) Project: Precision medicine for psychiatry. Am J Psychiatry 171:395–397.

2. Kapur S, Phillips AG, Insel TR (2012): Why has it taken so long for biological psychiatry to develop clinical tests and what to do about it? Mol Psychiatry 17:1174–1179.
3. Cuthbert BN, Insel TR (2013): Toward the future of psychiatric diagnosis: The seven pillars of RDoC. BMC Med 11:126.
4. Marquand AF, Rezek I, Buitelaar J, Beckmann CF (2016): Understanding heterogeneity in clinical cohorts using normative models: Beyond case-control studies. Biol Psychiatry 80:552–561.
5. Bzdok D, Meyer-Lindenberg A (2018): Machine learning for precision psychiatry: Opportunities and challenges. Biol Psychiatry Cogn Neurosci Neuroimaging 3:223–230.
6. Stephan KE, Bach DR, Fletcher PC, Flint J, Frank MJ, Friston KJ, *et al.* (2016): Charting the landscape of priority problems in psychiatry, part 1: Classification and diagnosis. Lancet Psychiatry 3:77–83.
7. Huys QJ, Maia TV, Frank MJ (2016): Computational psychiatry as a bridge from neuroscience to clinical applications. Nat Neurosci 19:404–413.
8. Wolfers T, Floris DL, Dinga R, van Rooij D, Isakoglou C, Kia SM, *et al.* (2019): From pattern classification to stratification: Towards conceptualizing the heterogeneity of autism spectrum disorder. Neurosci Biobehav Rev 104:240–254.
9. Lord C, Brugha TS, Charman T, Cusack J, Dumas G, Frazier T, *et al.* (2020): Autism spectrum disorder. Nat Rev Dis Primers 6:5.
10. Lombardo MV, Lai M-C, Baron-Cohen S (2019): Big data approaches to decomposing heterogeneity across the autism spectrum. Mol Psychiatr 24:1435–1450.
11. Grzadzinski R, Huerta M, Lord C (2013): DSM-5 and autism spectrum disorders (ASDs): An opportunity for identifying ASD subtypes. Mol Autism 4:12.
12. Jeste SS, Geschwind DH (2014): Disentangling the heterogeneity of autism spectrum disorder through genetic findings. Nat Rev Neurol 10:74–81.
13. Hansen SN, Schendel DE, Parner ET (2015): Explaining the increase in the prevalence of autism spectrum disorders: The proportion attributable to changes in reporting practices. JAMA Pediatr 169:56–62.
14. Buckley PF, Miller BJ, Lehrer DS, Castle DJ (2009): Psychiatric comorbidities and schizophrenia. Schizophr Bull 35:383–402.
15. Spoorthy MS, Chakrabarti S, Grover S (2019): Comorbidity of bipolar and anxiety disorders: An overview of trends in research. World J Psychiatry 9:7–29.
16. Ozonoff S, Heung K, Byrd R, Hansen R, Hertz-Picciotto I (2008): The onset of autism: Patterns of symptom emergence in the first years of life. Autism Res 1:320–328.
17. Pickles A, Anderson DK, Lord C (2014): Heterogeneity and plasticity in the development of language: A 17-year follow-up of children referred early for possible autism. J Child Psychol Psychiatry 55:1354–1362.
18. Charman T, Pickles A, Simonoff E, Chandler S, Loucas T, Baird G (2011): IQ in children with autism spectrum disorders: Data from the Special Needs and Autism Project (SNAP). Psychol Med 41:619–627.
19. Vargason T, Frye RE, McGuinness DL, Hahn J (2019): Clustering of co-occurring conditions in autism spectrum disorder during early childhood: A retrospective analysis of medical claims data. Autism Res 12:1272–1285.
20. Asperger H (1944): Die "Autistischen psychopathen" im kindesalter. Archiv für psychiatrie und nervenkrankheiten 117:76–136.
21. Kanner L (1943): Autistic disturbances of affective contact. Nerv Child 2:217–250.
22. Boat TF, Wu JT, Committee to Evaluate the Supplemental Security Income Disability Program for Children with Mental Disorders, Board on the Health of Select Populations, Board on Children, Youth, and Families, Institute of Medicine, *et al.* (2015): Prevalence of autism spectrum disorder. In: Boat TF, Wu JT, editors. Mental Disorders and Disabilities Among Low-Income Children. Washington, DC: National Academies Press, pp 241–266.
23. Christensen DL, Braun KVN, Baio J, Bilder D, Charles J, Constantino JN, *et al.* (2018): Prevalence and characteristics of autism spectrum disorder among children aged 8 years—Autism and Developmental Disabilities Monitoring Network, 11 sites, United States, 2012. MMWR Surveill Summ 65:1–23.
24. Rødgaard E-M, Jensen K, Vergnes J-N, Soulières I, Mottron L (2019): Temporal changes in effect sizes of studies comparing individuals with and without autism: A meta-analysis. JAMA Psychiatry 76:1124–1132.
25. Szatmari P, Georgiades S, Duku E, Bennett TA, Bryson S, Fombonne E, *et al.* (2015): Developmental trajectories of symptom severity and adaptive functioning in an inception cohort of preschool children with autism spectrum disorder. JAMA Psychiatry 72:276–283.
26. Fountain C, Winter AS, Bearman PS (2012): Six developmental trajectories characterize children with autism. Pediatrics 129:e1112–e1120.
27. Gotham K, Pickles A, Lord C (2012): Trajectories of autism severity in children using standardized ADOS scores. Pediatrics 130:e1278–e1284.
28. Kim SH, Macari S, Koller J, Chawarska K (2016): Examining the phenotypic heterogeneity of early autism spectrum disorder: Subtypes and short-term outcomes. J Child Psychol Psychiatry 57:93–102.
29. Lord C, Bishop S, Anderson D (2015): Developmental trajectories as autism phenotypes. Am J Med Genet C Semin Med Genet 169:198–208.
30. Magiati I, Tay XW, Howlin P (2014): Cognitive, language, social and behavioural outcomes in adults with autism spectrum disorders: A systematic review of longitudinal follow-up studies in adulthood. Clin Psychol Rev 34:73–86.
31. Bacon EC, Dufek S, Schreibman L, Stahmer AC, Pierce K, Courchesne E (2014): Measuring outcome in an early intervention program for toddlers with autism spectrum disorder: Use of a curriculum-based assessment. Autism Res Treat 2014:9.
32. Sestan N, State MW (2018): Lost in translation: Traversing the complex path from genomics to therapeutics in autism spectrum disorder. Neuron 100:406–423.
33. de la Torre-Ubieta L, Won H, Stein JL, Geschwind DH (2016): Advancing the understanding of autism disease mechanisms through genetics. Nat Med 22:345–361.
34. Bilder RM, Sabb FW, Cannon TD, London ED, Jentsch JD, Parker DS, *et al.* (2009): Phenomics: The systematic study of phenotypes on a genome-wide scale. Neuroscience 164:30–42.
35. Van Dam NT, O'Connor D, Marcelle ET, Ho EJ, Cameron Craddock R, Tobe RH, *et al.* (2017): Data-driven phenotypic categorization for neurobiological analyses: Beyond DSM-5 labels. Biol Psychiatry 81:484–494.
36. Haufe S, Meinecke F, Gorgen K, Dahne S, Haynes JD, Blankertz B, *et al.* (2014): On the interpretation of weight vectors of linear models in multivariate neuroimaging. NeuroImage 87:96–110.
37. Varoquaux G, Thirion B (2014): How machine learning is shaping cognitive neuroimaging. Gigascience 3:28.
38. Jimura K, Poldrack RA (2012): Analyses of regional-average activation and multivoxel pattern information tell complementary stories. Neuropsychologia 50:544–552.
39. Wing L, Gould J (1979): Severe impairments of social interaction and associated abnormalities in children: Epidemiology and classification. J Autism Dev Disord 9:11–29.
40. Lord C, Jones RM (2012): Annual research review: Re-thinking the classification of autism spectrum disorders. J Child Psychol Psychiatry 53:490–509.
41. American Psychiatric Association (2013): Diagnostic and Statistical Manual of Mental Disorders (DSM-5). Washington, DC: American Psychiatric Association Publishing.
42. World Health Organization (2019): International Statistical Classification of Diseases and Related Health Problems, 11th ed. Geneva, Switzerland: World Health Organization.
43. Tripoliti EE, Papadopoulos TG, Karanasiou GS, Naka KK, Fotiadis DI (2017): Heart failure: Diagnosis, severity estimation and prediction of adverse events through machine learning techniques. Comput Struct Biotechnol J 15:26–47.
44. Sandin S, Lichtenstein P, Kuja-Halkola R, Hultman C, Larsson H, Reichenberg A (2017): The heritability of autism spectrum disorder. JAMA 318:1182–1184.

45. Grove J, Ripke S, Als TD, Mattheisen M, Walters RK, Won H, *et al.* (2019): Identification of common genetic risk variants for autism spectrum disorder. Nat Genet 51:431–444.
46. Willsey AJ, State MW (2015): Autism spectrum disorders: From genes to neurobiology. Curr Opin Neurobiol 30:92–99.
47. SPARK Consortium (2018): SPARK: A US cohort of 50,000 families to accelerate autism research. Neuron 97:488–493.
48. Sanders SJ, He X, Willsey AJ, Ercan-Sencicek AG, Samocha KE, Cicek AE, *et al.* (2015): Insights into autism spectrum disorder genomic architecture and biology from 71 risk loci. Neuron 87:1215–1233.
49. Stessman HA, Bernier R, Eichler EE (2014): A genotype-first approach to defining the subtypes of a complex disease. Cell 156:872–877.
50. Simons VIP Consortium (2012): Simons Variation in Individuals Project (Simons VIP): A genetics-first approach to studying autism spectrum and related neurodevelopmental disorders. Neuron 73:1063–1067.
51. Bishop SL, Farmer C, Bal V, Robinson EB, Willsey AJ, Werling DM, *et al.* (2017): Identification of developmental and behavioral markers associated with genetic abnormalities in autism spectrum disorder. Am J Psychiatry 174:576–585.
52. Abrahams BS, Geschwind DH (2008): Advances in autism genetics: On the threshold of a new neurobiology. Nat Rev Genet 9:341–355.
53. Feczko E, Balba NM, Miranda-Dominguez O, Cordova M, Karalunas SL, Irwin L, *et al.* (2018): Subtyping cognitive profiles in autism spectrum disorder using a functional random forest algorithm. NeuroImage 172:674–688.
54. Van Rooij D, Anagnostou E, Arango C, Auzias G, Behrmann M, Busatto GF, *et al.* (2018): Cortical and subcortical brain morphometry differences between patients with autism spectrum disorder and healthy individuals across the lifespan: Results from the ENIGMA ASD Working Group. Am J Psychiatry 175:359–369.
55. Bedford SA, Park MTM, Devenyi GA, Tullo S, Germann J, Patel R, *et al.* (2020): Large-scale analyses of the relationship between sex, age and intelligence quotient heterogeneity and cortical morphometry in autism spectrum disorder. Mol Psychiatr 25:614–628.
56. Picci G, Gotts SJ, Scherf KS (2016): A theoretical rut: Revisiting and critically evaluating the generalized under/over-connectivity hypothesis of autism. Dev Sci 19:524–549.
57. Hernandez LM, Rudie JD, Green SA, Bookheimer S, Dapretto M (2015): Neural signatures of autism spectrum disorders: Insights into brain network dynamics. Neuropsychopharmacology 40:171–189.
58. Bernhardt BC, Di Martino A, Valk SL, Wallace GL (2017): Neuroimaging-based phenotyping of the autism ppectrum. Curr Top Behav Neurosci 30:341–355.
59. Abraham A, Milham MP, Di Martino A, Craddock RC, Samaras D, Thirion B, *et al.* (2017): Deriving reproducible biomarkers from multi-site resting-state data: An autism-based example. NeuroImage 147:736–745.
60. He Y, Byrge L, Kennedy DP (2020): Nonreplication of functional connectivity differences in autism spectrum disorder across multiple sites and denoising strategies. Hum Brain Mapp 41:1334–1350.
61. King JB, Prigge MBD, King CK, Morgan J, Weathersby F, Fox JC, *et al.* (2019): Generalizability and reproducibility of functional connectivity in autism. Mol Autism 10:27.
62. Holiga S, Hipp JF, Chatham CH, Garces P, Spooren W, D'Ardhuy XL, *et al.* (2019): Patients with autism spectrum disorders display reproducible functional connectivity alterations. Sci Transl Med 11:eaat9223.
63. Yahata N, Morimoto J, Hashimoto R, Lisi G, Shibata K, Kawakubo Y, *et al.* (2016): A small number of abnormal brain connections predicts adult autism spectrum disorder. Nat Commun 7:11254.
64. Easson AK, Fatima Z, McIntosh AR (2018): Functional connectivity-based subtypes of individuals with and without autism spectrum disorder. Netw Neurosci 3:344–362.
65. Duffy FH, Als H (2019): Autism, spectrum or clusters? An EEG coherence study. BMC Neurol 19:27.
66. Chen H, Uddin LQ, Guo X, Wang J, Wang R, Wang X, *et al.* (2019): Parsing brain structural heterogeneity in males with autism spectrum disorder reveals distinct clinical subtypes. Hum Brain Mapp 40:628–637.
67. Jao Keehn RJ, Nair S, Pueschel EB, Linke AC, Fishman I, Muller RA (2019): Atypical local and distal patterns of occipito-frontal functional connectivity are related to symptom severity in autism. Cereb Cortex 29:3319–3330.
68. Hrdlicka M, Dudova I, Beranova I, Lisy J, Belsan T, Neuwirth J, *et al.* (2005): Subtypes of autism by cluster analysis based on structural MRI data. Eur Child Adolesc Psychiatry 14:138–144.
69. Hong S-J, Valk SL, Di Martino A, Milham MP, Bernhardt BC (2018): Multidimensional neuroanatomical subtyping of autism spectrum disorder. Cereb Cortex 28:3578–3588.
70. Bethlehem RAI, Seidlitz J, Romero-Garcia R, Lombardo MV (2018): Using normative age modelling to isolate subsets of individuals with autism expressing highly age-atypical cortical thickness features. bioRxiv. https://doi.org/10.1101/252593.
71. Stefanik L, Erdman L, Ameis SH, Foussias G, Mulsant BH, Behdinan T, *et al.* (2018): Brain-behavior participant similarity networks among youth and emerging adults with schizophrenia spectrum, autism spectrum, or bipolar disorder and matched controls. Neuropsychopharmacology 43:1180–1188.
72. Kernbach JM, Satterthwaite TD, Bassett DS, Smallwood J, Margulies D, Krall S, *et al.* (2018): Shared endo-phenotypes of default mode dysfunction in attention deficit/hyperactivity disorder and autism spectrum disorder. Transl Psychiatry 8:133.
73. Lefebvre A, Delorme R, Delanoe C, Amsellem F, Beggiato A, Germanaud D, *et al.* (2018): Alpha waves as a neuromarker of autism spectrum disorder: The challenge of reproducibility and heterogeneity. Front Neurosci 12:662.
74. Zabihi M, Oldehinkel M, Wolfers T, Frouin V, Goyard D, Loth E, *et al.* (2019): Dissecting the heterogeneous cortical anatomy of autism spectrum disorder using normative models. Biol Psychiatry Cogn Neurosci Neuroimaging 4:567–578.
75. Tang S, Sun N, Floris DL, Zhang X, Di Martino A, Yeo BTT (2019): Reconciling dimensional and categorical models of autism heterogeneity: A brain connectomics and behavioral study [published online ahead of print Nov 18]. Biol Psychiatry.
76. Krueger RF, Eaton NR (2015): Transdiagnostic factors of mental disorders. World Psychiatry 14:27–29.
77. Milham MP, Craddock RC, Son JJ, Fleischmann M, Clucas J, Xu H, *et al.* (2018): Assessment of the impact of shared brain imaging data on the scientific literature. Nat Commun 9:2818.
78. Di Martino A, Yan CG, Li Q, Denio E, Castellanos FX, Alaerts K, *et al.* (2014): The Autism Brain Imaging Data Exchange: Towards a large-scale evaluation of the intrinsic brain architecture in autism. Mol Psychiatry 19:659–667.
79. Di Martino A, O'Connor D, Chen B, Alaerts K, Anderson JS, Assaf M, *et al.* (2017): Enhancing studies of the connectome in autism using the Autism Brain Imaging Data Exchange II. Sci Data 4:170010.
80. Loth E, Charman T, Mason L, Tillmann J, Jones EJH, Wooldridge C, *et al.* (2017): The EU-AIMS Longitudinal European Autism Project (LEAP): Design and methodologies to identify and validate stratification biomarkers for autism spectrum disorders. Mol Autism 8:24.
81. Kooijman MN, Kruithof CJ, van Duijn CM, Duijts L, Franco OH, van IMH, *et al.* (2016): The Generation R Study: Design and cohort update 2017. Eur J Epidemiol 31:1243–1264.
82. Hall D, Huerta MF, McAuliffe MJ, Farber GK (2012): Sharing heterogeneous data: The National Database for Autism Research. Neuroinformatics 10:331–339.
83. Nooner KB, Colcombe SJ, Tobe RH, Mennes M, Benedict MM, Moreno AL, *et al.* (2012): The NKI-Rockland sample: A model for accelerating the pace of discovery science in psychiatry. Front Neurosci 6:152.
84. Al-Jawahiri R, Milne E (2017): Resources available for autism research in the big data era: A systematic review. PeerJ 5, e2880.

85. Casanova MF, El-Baz AS, Kamat SS, Dombroski BA, Khalifa F, Elnakib A, *et al.* (2013): Focal cortical dysplasias in autism spectrum disorders. Acta Neuropathol Commun 1:67.
86. Avino TA, Hutsler JJ (2010): Abnormal cell patterning at the cortical gray-white matter boundary in autism spectrum disorders. Brain Res 1360:138–146.
87. Wegiel J, Kuchna I, Nowicki K, Imaki H, Wegiel J, Marchi E, *et al.* (2010): The neuropathology of autism: Defects of neurogenesis and neuronal migration, and dysplastic changes. Acta Neuropathol 119:755–770.
88. Saeys Y, Inza I, Larranaga P (2007): A review of feature selection techniques in bioinformatics. Bioinformatics 23:2507–2517.
89. Dy JG, Brodley CE (2004): Feature selection for unsupervised learning. J Mach Learn Res 5:845–889.
90. Leek JT, Storey JD (2007): Capturing heterogeneity in gene expression studies by surrogate variable analysis. PLoS Genet 3:1724–1735.
91. Clarke R, Ressom HW, Wang A, Xuan J, Liu MC, Gehan EA, *et al.* (2008): The properties of high-dimensional data spaces: Implications for exploring gene and protein expression data. Nat Rev Cancer 8:37–49.
92. Hughes G (1968): On the mean accuracy of statistical pattern recognizers. IEEE Trans Inform Theory 14:55–63.
93. Wang B, Mezlini AM, Demir F, Fiume M, Tu Z, Brudno M, *et al.* (2014): Similarity network fusion for aggregating data types on a genomic scale. Nat Methods 11:333–337.
94. Shen C, Vogelstein JT, Priebe CE (2017): Manifold matching using shortest-path distance and joint neighborhood selection. Pattern Recognit Lett 92:41–48.
95. Pai S, Bader GD (2018): Patient similarity networks for precision medicine. J Mol Biol 430:2924–2938.
96. Ruan P, Wang Y, Shen R, Wang S (2019): Using association signal annotations to boost similarity network fusion. Bioinformatics 35:3718–3726.
97. Feczko E, Miranda-Dominguez O, Marr M, Graham AM, Nigg JT, Fair DA (2019): The heterogeneity problem: Approaches to identify psychiatric subtypes. Trends Cogn Sci 23:584–601.
98. Xu C, Tao D, Xu C (2013): A survey on multi-view learning. arXiv: 1304–5634.
99. Lock EF, Hoadley KA, Marron JS, Nobel AB (2013): Joint and individual variation explained (JIVE) for integrated analysis of multiple data types. Ann Appl Stat 7:523–542.
100. Zhao Y, Klein A, Castellanos FX, Milham MP (2019): Brain age prediction: Cortical and subcortical shape covariation in the developing human brain. NeuroImage 202:116149.
101. Smith SM, Nichols TE, Vidaurre D, Winkler AM, Behrens TE, Glasser MF, *et al.* (2015): A positive-negative mode of population covariation links brain connectivity, demographics and behavior. Nat Neurosci 18:1565–1567.
102. Drysdale AT, Grosenick L, Downar J, Dunlop K, Mansouri F, Meng Y, *et al.* (2017): Resting-state connectivity biomarkers define neurophysiological subtypes of depression. Nat Med 23:28–38.
103. Kaufman S, Rosset S, Perlich C, Stitelman O (2012): Leakage in data mining: Formulation, detection, and avoidance. TKDD 6:15.
104. Castellanos FX, Di Martino A, Craddock RC, Mehta AD, Milham MP (2013): Clinical applications of the functional connectome. NeuroImage 80:527–540.
105. Makowski C, Lepage M, Evans AC (2019): Head motion: The dirty little secret of neuroimaging in psychiatry. J Psychiatry Neurosci 44:62–68.
106. Power JD, Barnes KA, Snyder AZ, Schlaggar BL, Petersen SE (2012): Spurious but systematic correlations in functional connectivity MRI networks arise from subject motion. NeuroImage 59:2142–2154.
107. Satterthwaite TD, Wolf DH, Loughead J, Ruparel K, Elliott MA, Hakonarson H, *et al.* (2012): Impact of in-scanner head motion on multiple measures of functional connectivity: Relevance for studies of neurodevelopment in youth. NeuroImage 60:623–632.
108. Yan CG, Cheung B, Kelly C, Colcombe S, Craddock RC, Di Martino A, *et al.* (2013): A comprehensive assessment of regional variation in the impact of head micromovements on functional connectomics. NeuroImage 76:183–201.
109. Parkes L, Fulcher B, Yucel M, Fornito A (2018): An evaluation of the efficacy, reliability, and sensitivity of motion correction strategies for resting-state functional MRI. NeuroImage 171:415–436.
110. Reuter M, Tisdall MD, Qureshi A, Buckner RL, van der Kouwe AJW, Fischl B (2015): Head motion during MRI acquisition reduces gray matter volume and thickness estimates. NeuroImage 107:107–115.
111. Yendiki A, Koldewyn K, Kakunoori S, Kanwisher N, Fischl B (2014): Spurious group differences due to head motion in a diffusion MRI study. NeuroImage 88:79–90.
112. Koldewyn K, Yendiki A, Weigelt S, Gweon H, Julian J, Richardson H, *et al.* (2014): Differences in the right inferior longitudinal fasciculus but no general disruption of white matter tracts in children with autism spectrum disorder. Proc Natl Acad Sci U S A 111:1981–1986.
113. Siegel JS, Power JD, Dubis JW, Vogel AC, Church JA, Schlaggar BL, *et al.* (2014): Statistical improvements in functional magnetic resonance imaging analyses produced by censoring high-motion data points. Hum Brain Mapp 35:1981–1996.
114. Dosenbach NUF, Koller JM, Earl EA, Miranda-Dominguez O, Klein RL, Van AN, *et al.* (2017): Real-time motion analytics during brain MRI improve data quality and reduce costs. NeuroImage 161:80–93.
115. Ai L, Craddock RC, Tottenham N, Dyke JP, Colcombe S, Milham M, *et al.* (2019): Is it time to switch your T1W sequence? Assessing the impact of prospective motion correction on the reliability and quality of structural imaging. bioRxiv. https://doi.org/10.1101/666289.
116. Pruim RHR, Mennes M, van Rooij D, Llera A, Buitelaar JK, Beckmann CF (2015): ICA-AROMA: A robust ICA-based strategy for removing motion artifacts from fMRI data. NeuroImage 112:267–277.
117. Zuo XN, Xu T, Milham MP (2019): Harnessing reliability for neuroscience research. Nat Hum Behav 3:768–771.
118. Iscan Z, Jin TB, Kendrick A, Szeglin B, Lu H, Trivedi M, *et al.* (2015): Test-retest reliability of FreeSurfer measurements within and between sites: Effects of visual approval process. Hum Brain Mapp 36:3472–3485.
119. Madan CR, Kensinger EA (2017): Test-retest reliability of brain morphology estimates. Brain Inform 4:107–121.
120. O'Connor D, Potler NV, Kovacs M, Xu T, Ai L, Pellman J, *et al.* (2017): The Healthy Brain Network Serial Scanning Initiative: A resource for evaluating inter-individual differences and their reliabilities across scan conditions and sessions. Gigascience 6:1–14.
121. Elliott ML, Knodt AR, Cooke M, Kim MJ, Melzer TR, Keenan R, *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 189:516–532.
122. Marquand AF, Wolfers T, Mennes M, Buitelaar J, Beckmann CF (2016): Beyond lumping and splitting: A review of computational approaches for stratifying psychiatric disorders. Biol Psychiatry Cogn Neurosci Neuroimaging 1:433–447.
123. Liu Y, Li Z, Xiong H, Gao X, Wu J (2010): Understanding of internal clustering validation measures. In: ICDM '10: Proceedings of the 2010 IEEE International Conference on Data Mining Red Hook. NY: Curran Associates, 911–916.
124. Priebe CE, Park Y, Vogelstein JT, Conroy JM, Lyzinski V, Tang M, *et al.* (2019): On a two-truths phenomenon in spectral graph clustering. Proc Natl Acad Sci U S A 116:5995–6000.
125. Zhang X, Mormino EC, Sun N, Sperling RA, Sabuncu MR, Yeo BT, *et al.* (2016): Bayesian model reveals latent atrophy factors with dissociable cognitive trajectories in Alzheimer's disease. Proc Natl Acad Sci U S A 113:E6535–E6544.
126. Thompson WK, Hallmayer J, O'Hara R, Alzheimer's Disease Neuroimaging Initiative (2011): Design considerations for characterizing psychiatric trajectories across the lifespan: Application to effects of APOE-ε4 on cerebral cortical thickness in Alzheimer's disease. Am J Psychiatry 168:894–903.
127. Dean DC, Dirks H, O'Muircheartaigh J, Walker L, Jerskey BA, Lehman K, *et al.* (2014): Pediatric neuroimaging using magnetic

resonance imaging during non-sedated sleep. Pediatr Radiol 44:64–72.
128. Manning JH, Courchesne E, Fox PT (2013): Intrinsic connectivity network mapping in young children during natural sleep. NeuroImage 83:288–293.
129. Bolton TA, Jochaut D, Giraud AL, Van De Ville D (2018): Brain dynamics in ASD during movie-watching show idiosyncratic functional integration and segregation. Hum Brain Mapp 39:2391–2404.
130. Vanderwal T, Kelly C, Eilbott J, Mayes LC, Castellanos FX (2015): Inscapes: A movie paradigm to improve compliance in functional magnetic resonance imaging. NeuroImage 122:222–232.
131. Nordahl CW, Mello M, Shen AM, Shen MD, Vismara LA, Li D, *et al.* (2016): Methods for acquiring MRI data in children with autism spectrum disorder and intellectual impairment without the use of sedation. J Neurodev Disord 8:20.
132. Tisdall MD, Reuter M, Qureshi A, Buckner RL, Fischl B, van der Kouwe AJW (2016): Prospective motion correction with volumetric navigators (vNavs) reduces the bias and variance in brain morphometry induced by subject motion. NeuroImage 127:11–22.
133. Loomes R, Hull L, Mandy WPL (2017): What is the male-to-female ratio in autism spectrum disorder? A systematic review and meta-analysis. J Am Acad Child Adolesc Psychiatry 56:466–474.
134. Lai MC, Lombardo MV, Auyeung B, Chakrabarti B, Baron-Cohen S (2015): Sex/gender differences and autism: Setting the scene for future research. J Am Acad Child Adolesc Psychiatry 54:11–24.
135. Floris DL, Lai MC, Nath T, Milham MP, Di Martino A (2018): Network-specific sex differentiation of intrinsic brain function in males with autism. Mol Autism 9:17.
136. Irimia A, Torgerson CM, Jacokes ZJ, Van Horn JD (2017): The connectomes of males and females with autism spectrum disorder have significantly different white matter connectivity densities. Sci Rep 7:46401.
137. Button KS, Ioannidis JP, Mokrysz C, Nosek BA, Flint J, Robinson ES, *et al.* (2013): Power failure: Why small sample size undermines the reliability of neuroscience. Nat Rev Neurosci 14:365–376.
138. Baker M (2016): 1,500 scientists lift the lid on reproducibility. Nature 533:452–454.
139. Breiman L (1996): Bagging predictors. Mach Learn 24:123–140.
140. Nikolaidis A, Solon Heinsfeld A, Xu T, Bellec P, Vogelstein J, Milham M (2020): Bagging improves reproducibility of functional parcellation of the human brain. NeuroImage 214:116678.
141. Kelly C, Toro R, Di Martino A, Cox CL, Bellec P, Castellanos FX, *et al.* (2012): A convergent functional architecture of the insula emerges across imaging modalities. NeuroImage 61:1129–1142.
142. Bellec P, Rosa-Neto P, Lyttelton OC, Benali H, Evans AC (2010): Multi-level bootstrap analysis of stable clusters in resting-state fMRI. NeuroImage 51:1126–1139.
143. Martinez-Murcia FJ, Lai MC, Gorriz JM, Ramirez J, Young AM, Deoni SC, *et al.* (2017): On the brain structure heterogeneity of autism: Parsing out acquisition site effects with significance-weighted principal component analysis. Hum Brain Mapp 38:1208–1223.
144. Leek JT, Johnson WE, Parker HS, Jaffe AE, Storey JD (2012): The sva package for removing batch effects and other unwanted variation in high-throughput experiments. Bioinformatics 28:882–883.
145. Nielson DM, Pereira F, Zheng CY, Migineishvili N, Lee JA, Thomas AG, *et al.* (2018): Detecting and harmonizing scanner differences in the ABCD Study—Annual release 1.0. bioRxiv. doi:10.1101/309260.
146. Herrington JD, Miller JS, Pandey J, Schultz RT (2016): Anxiety and social deficits have distinct relationships with amygdala function in autism spectrum disorder. Soc Cogn Affect Neurosci 11:907–914.
147. Bennett RH, Somandepalli K, Roy AK, Di Martino A (2017): The neural correlates of emotional lability in children with autism spectrum disorder. Brain Connect 7:281–288.
148. Chantiluke K, Christakou A, Murphy CM, Giampietro V, Daly EM, Ecker C, *et al.* (2014): Disorder-specific functional abnormalities during temporal discounting in youth with attention deficit hyperactivity disorder (ADHD), autism and comorbid ADHD and autism. Psychiatry Res 223:113–120.
149. Di Martino A, Zuo X-N, Kelly C, Grzadzinski R, Mennes M, Schvarcz A, *et al.* (2013): Shared and distinct intrinsic functional network centrality in autism and attention-deficit/hyperactivity disorder. Biol Psychiatry 74:623–632.
150. Courchesne E, Pierce K (2005): Why the frontal cortex in autism might be talking only to itself: Local over-connectivity but long-distance disconnection. Curr Opin Neurobiol 15:225–230.
151. Smith RX, Jann K, Dapretto M, Wang DJJ (2018): Imbalance of functional connectivity and temporal entropy in resting-state networks in autism spectrum disorder: A machine learning approach. Front Neurosci 12:869.
152. Hahamy A, Behrmann M, Malach R (2015): The idiosyncratic brain: Distortion of spontaneous connectivity patterns in autism spectrum disorder. Nat Neurosci 18:302–309.
153. Hong SJ, de Wael RV, Bethlehem RAI, Lariviere S, Paquola C, Valk SL, *et al.* (2019): Atypical functional connectome hierarchy in autism. Nat Commun 10:1022.
154. Willsey AJ, Sanders SJ, Li M, Dong S, Tebbenkamp AT, Muhle RA, *et al.* (2013): Coexpression networks implicate human midfetal deep cortical projection neurons in the pathogenesis of autism. Cell 155:997–1007.
155. Geschwind DH, State MW (2015): Gene hunting in autism spectrum disorder: On the path to precision medicine. Lancet Neurol 14:1109–1120.
156. Zhou Y, Sharma J, Ke Q, Landman R, Yuan J, Chen H, *et al.* (2019): Atypical behaviour and connectivity in SHANK3-mutant macaques. Nature 570:326–331.
157. Liska A, Gozzi A (2016): Can mouse imaging studies bring order to autism connectivity chaos? Front Neurosci 10:484.
158. Gozzi A, Schwarz AJ (2016): Large-scale functional connectivity networks in the rodent brain. NeuroImage 127:496–509.
159. Bertero A, Liska A, Pagani M, Parolisi R, Masferrer ME, Gritti M, *et al.* (2018): Autism-associated 16p11.2 microdeletion impairs prefrontal functional connectivity in mouse and human. Brain 141:2055–2065.
160. Ellegood J, Anagnostou E, Babineau BA, Crawley JN, Lin L, Genestine M, *et al.* (2015): Clustering autism: Using neuroanatomical differences in 26 mouse models to gain insight into the heterogeneity. Mol Psychiatry 20:118–125.
161. Liska A, Galbusera A, Schwarz AJ, Gozzi A (2015): Functional connectivity hubs of the mouse brain. NeuroImage 115:281–291.
162. Bertero A, Liska A, David G, Galbusera A, Pasqualetti M, Gozzi A (2015): Frontal hypoconnectivity in the 16p11.2 microdeletion autism model. In: Presented at the annual meeting of the Society for Neuroscience, October 17–21, Chicago, Illinois.
163. Moore A, Wozniak M, Yousef A, Barnes CC, Cha D, Courchesne E, *et al.* (2018): The geometric preference subtype in ASD: Identifying a consistent, early-emerging phenomenon through eye tracking. Mol Autism 9:19.
164. Lombardo MV, Eyler L, Moore A, Datko M, Carter Barnes C, Cha D, *et al.* (2019): Default mode-visual network hypoconnectivity in an autism subtype with pronounced social visual engagement difficulties. eLife 8:e47427.
165. Pierce K, Marinero S, Hazin R, McKenna B, Barnes CC, Malige A (2016): Eye tracking reveals abnormal visual preference for geometric images as an early biomarker of an autism spectrum disorder subtype associated with increased symptom severity. Biol Psychiatry 79:657–666.
166. Pierce K, Conant D, Hazin R, Stoner R, Desmond J (2011): Preference for geometric patterns early in life as a risk factor for autism. Arch Gen Psychiatry 68:101–109.
167. Webb SJ, Shic F, Murias M, Sugar CA, Naples AJ, Barney E, *et al.* (2019): Biomarker acquisition and quality control for multi-site studies: The Autism Biomarkers Consortium for Clinical Trials. Front Integr Neurosci 13:71.
168. Kang E, Keifer CM, Levy EJ, Foss-Feig JH, McPartland JC, Lerner MD (2018): Atypicality of the N170 event-related potential in

autism spectrum disorder: A meta-analysis. Biol Psychiatry Cogn Neurosci Neuroimaging 3:657–666.

169. Cohen IL, Gardner JM, Karmel BZ, Kim S-Y (2014): Rating scale measures are associated with Noldus EthoVision-XT video tracking of behaviors of children on the autism spectrum. Mol Autism 5:15.

170. Jones EJ, Venema K, Earl R, Lowy R, Barnes K, Estes A, *et al.* (2016): Reduced engagement with social stimuli in 6-month-old infants with later autism spectrum disorder: A longitudinal prospective study of infants at high familial risk. J Neurodev Disord 8:7.

171. Dosenbach NU, Nardos B, Cohen AL, Fair DA, Power JD, Church JA, *et al.* (2010): Prediction of individual brain maturity using fMRI. Science 329:1358–1361.

172. Schaefer A, Kong R, Gordon EM, Laumann TO, Zuo XN, Holmes AJ, *et al.* (2018): Local-global parcellation of the human cerebral cortex from intrinsic functional connectivity MRI. Cereb Cortex 28:3095–3114.

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

174. Ameis SH, Lerch JP, Taylor MJ, Lee W, Viviano JD, Pipitone J, *et al.* (2016): A diffusion tensor imaging study in children with ADHD, autism spectrum disorder, OCD, and matched controls: Distinct and non-distinct white matter disruption and dimensional brain-behavior relationships. Am J Psychiatry 173:1213–1222.