



RESEARCH ARTICLE

Exclusion of females in autism research: Empirical evidence for a “leaky” recruitment-to-research pipeline

Anila M. D'Mello¹  | Isabelle R. Frosch²  | Cindy E. Li³ | Annie L. Cardinaux^{1,4} | John D.E. Gabrieli^{1,4}

¹McGovern Institute for Brain Research, Massachusetts Institute of Technology, Cambridge, Massachusetts, USA

²Department of Psychology, Northwestern University, Evanston, Illinois, USA

³Hock E. Tan and K. Lisa Yang Center for Autism Research at Massachusetts Institute of Technology, Cambridge, Massachusetts, USA

⁴Department of Brain and Cognitive Sciences, Massachusetts Institute of Technology, Cambridge, Massachusetts, USA

Correspondence

Anila M. D'Mello, McGovern Institute for Brain Research, Massachusetts Institute of Technology, 43 Vassar Street, Cambridge, MA 02139, USA.

Email: admello@mit.edu

Present address

Anila M. D'Mello, Department of Psychiatry, University of Texas Southwestern Medical Center, Dallas, TX; Peter J. O'Donnell Brain Institute, University of Texas Southwestern Medical Center, Dallas, TX; Department of Psychology, University of Texas at Dallas, Cambridge, Richardson, USA.

Funding information

Hock E. Tan and K. Lisa Yang Center for Autism Research; National Institute of Mental Health; Simons Center for the Social Brain at MIT; National Institutes of Mental Health, Grant/Award Number: F32 MH117933

Abstract

Autism spectrum disorder (ASD) is characterized by challenges in social communication and the presence of repetitive behaviors or restricted interests. Notably, males are four times as likely as females to be diagnosed with autism. Despite efforts to increase representation and characterization of autistic females, research studies consistently enroll small samples of females, or exclude females altogether. Importantly, researchers often rely on standardized measures to confirm diagnosis prior to enrollment in research studies. We retrospectively analyzed the effects of one such measure (Autism Diagnostic Observation Schedule, ADOS) on research inclusion/exclusion rates by sex in autistic adults, all of whom had a preexisting community diagnosis of autism ($n = 145$, 95 male, 50 female). Using the ADOS as a confirmatory diagnostic measure resulted in the exclusion of autistic females at a rate over 2.5 times higher than that of autistic males. We compared sex ratios in our sample to those in other large, publically available datasets that rely either on community diagnosis (6 datasets, total $n = 42,209$) or standardized assessments (2 datasets, total $n = 214$) to determine eligibility of participants for research. Reliance on community diagnosis rather than confirmatory diagnostic assessments resulted in significantly more equal sex ratios. These results provide evidence for a “leaky” recruitment-to-research pipeline for females in autism research.

Lay Summary

Despite efforts to increase the representation of autistic females in research, studies consistently enroll small samples of females or exclude females altogether. We find that despite making up almost 50% of the initially recruited sample based upon self-report of community diagnosis, autistic females are disproportionately excluded from research participation as a result of commonly used autism diagnostic measures. In our sample, and several other publically available datasets, reliance on community diagnosis resulted in significantly more equal sex ratios.

KEYWORDS

ABIDE, ADOS, Autism Physical Health Survey, autism spectrum disorder, Channel 4, diagnosis, exclusion criteria, females, IMAGES, inclusion criteria, LifeLines, Muscial Universe, recruitment, sex differences, SPARK

INTRODUCTION

Autism spectrum disorder (ASD) is a neurodevelopmental disorder characterized by challenges in social

communication and the presence of repetitive behaviors or restricted interests. Historically, autism has been viewed as a predominantly male disorder with male to female sex ratios typically reported as 4:1 (Lai,

Lombardo, et al., 2015) and ranging as high as 7:1 (Rutherford et al., 2016). Subsequent findings, however, suggest that the sex ratio discrepancy may be smaller than originally thought (Barnard-Brak et al., 2019), and even equal in some samples (see Burrows et al., 2022). In fact, ratios as low as 3:1 have been reported in children (Loomes et al., 2017) and may be even lower among adults (Posserud et al., 2021). In light of these findings, there have been calls to improve characterization of females on the autism spectrum and to increase their representation in research (Barnard-Brak et al., 2019; Jack et al., 2021; Watkins et al., 2014). Underrepresentation of women in research is not unique to autism. Across both basic research studies and clinical trials, males are often disproportionately overrepresented. This can lead to failures in the diagnosis and treatment of women, including reduced efficacy of and unforeseen negative side effects from pharmacological treatment (Shansky & Murphy, 2021), increased service needs and barriers to treatment relative to males, and other unmet treatment needs (Koffer Miller et al., 2022). Despite increased awareness of the underrepresentation of females in prevalence estimates and research studies, calls from the scientific and autism community to include more females in research, requirements of federal funding agencies, and the best efforts of researchers to recruit more females, empirical autism research studies continue to report small sample sizes of females or male-only studies. This poses a problem for both basic science and the clinical relevance of this research for females on the autism spectrum.

Of note, other neurodevelopmental conditions, such as attention-deficit/hyperactivity disorder (ADHD), provide strong evidence that females may be *overlooked* rather than absent. For example, ADHD girls are underdiagnosed relative to boys, and sex differences in prevalence can be attributed to clinician tendency to overlook less overt, yet clinically significant, symptoms in ADHD girls. This may be due in part to sex differences in symptom presentation (e.g., ADHD females are more likely to be inattentive than hyperactive) and to the perception that ADHD is rare in females (Hinshaw et al., 2022). Similarly, autism is perceived to be a primarily male disorder and is diagnosed less often in females despite equal or higher symptom severity (Cola et al., 2022; Lockwood Estrin et al., 2020; Mandy et al., 2012; Rea et al., 2022). In addition, autistic females may mask or camouflage symptoms (Lai et al., 2017), are sometimes misdiagnosed due to concurrent diagnoses (Lai, Baron-Cohen, & Buxbaum, 2015; Rutherford et al., 2016), have lower probabilities of referral for diagnosis (Cumin et al., 2021), and as a result tend to be diagnosed later in life (Fusar-Poli et al., 2020; Hiller et al., 2016). (A note on terminology: 1) We use identity first-language ["autistic"] in concordance with the expressed preference of many individuals on the autism spectrum; Bottema-Beutel et al., 2021. 2) We use binary sex-based terms "male" and "female" rather than gender-based terms due to the nature of our data set.) These findings of sex-based

factors in diagnostic outcomes provide a compelling framework for understanding underrepresentation of females in autism research.

Clinician and societal-level perceptual biases about autism coupled with actual sex-based phenotypic differences also have downstream effects on the development, scoring, and interpretation of standardized measures used to diagnose autism in the clinic and to confirm diagnosis in research (Rea et al., 2022). In research, gold-standard diagnostic measures (most commonly the Autism Diagnostic Observation Schedule, ADOS or the autism diagnostic interview, ADI) are often used to confirm diagnosis and to exclude participants who do not meet measure-specified cut-offs for autism. These tools are used even when a participant has received a prior community diagnosis (a diagnosis made by general medical practitioners, neuropsychologists, and mental health providers). Further, these measures are often used as the *sole* metric of inclusion with strict adherence to cut-offs and without the addition of clinical judgment. This not only diverges from what is recommended by the tools' creators (Lord et al., 2000), but is also quite different from community diagnostic practices. Community diagnostic tools are typically varied and can include behavioral assessment, clinical interviewing, parent report, medical and symptom history, and self-reports. In addition, community diagnosis largely relies on standard, empirically-derived criteria for ASD as outlined in the DSM-5 (and its international analogue: International Classification of Diseases, 11th Review; World Health Organization, 1992) rather than measure-specific cut-off scores. These practical and conceptual differences allow for potential discrepancies between community diagnosis and diagnosis as confirmed by the ADOS. In research, these differences coupled with evidence that females often score lower on the ADOS (and other confirmatory diagnostic measures) (e.g., Bastiaansen et al., 2011; Lai et al., 2011; Ratto et al., 2018) may result in increased exclusion of females from research participation.

In the current study, we analyzed data from individuals with an existing community diagnosis of ASD who were recruited to participate in autism research. We retrospectively examined sex-based differences in symptom severity as measured by the ADOS, and then investigated how using the ADOS as a confirmatory diagnostic assessment differentially affected inclusion rates of males and females in autism research. We then compared our site-specific findings to patterns of inclusion/exclusion by sex in several publically available databases that used either community diagnosis or confirmatory diagnostic assessments to determine inclusion in research.

METHODS

Participants

We conducted a retrospective review of an internal research database of autistic participants to determine

how application of confirmatory diagnostic assessments affected inclusion of males and females. The Autism Research Participant Database at the Massachusetts Institute of Technology (MIT) is a shared resource, established in 2007, and supported by the Simons Center for the Social Brain and the Hock E. Tan and K. Lisa Yang Center for Autism Research at MIT. The database consists of a total of 376 adults and children (291 males; 85 females) recruited from the community with an existing clinical diagnosis of autism, autism spectrum disorder, Asperger's syndrome, or PDD-NOS (prior to new criteria in the DSM-V) as reported by the participant. Participants were recruited to the database in an ongoing manner between 2007 and 2020, and matched with specific research studies based on participant interest and individual study criteria. To target the present analyses to verbally fluent adolescents and adults, we included only participants who were administered a Module 4 ADOS or ADOS-2 (best suited for verbally fluent adolescents and adults) and for whom all item-level ADOS scores were available. Of this group of participants, we further included only participants 16 years or older to reflect the ADOS manual recommendations regarding appropriate age ranges for Module 4 administration. This resulted in a final sample size of $n = 145$ individuals (95 males, 50 females; mean age = 29.55 ± 10.80 , ranging in age from 16 to 65 years old). Participant-reported information was used to categorize participants into one of two groups (male or female) that we refer to as "sex" in this manuscript. Given the retrospective nature of this study, we were limited by binary sex or gender options, depending on how questionnaires were worded. All participants provided informed consent to be included in the database (or caregiver consent and assent for minors).

Confirmatory diagnostic assessment process

Prior to participation and matching with available studies within MIT, research-reliable examiners (all female) administered the ADOS or ADOS-2: a standardized, semi-structured measure widely used in the assessment of autism (Lord, 2012; Lord et al., 2000). All participants in the current sample were administered the ADOS Module-4. ADOS scores and cut-offs are determined by an algorithm that combines scores across symptom domains. Higher ADOS scores are associated with greater frequency or degree of autism symptoms (Lord et al., 2000). Individuals who scored below the ADOS cut-off for "autism spectrum disorder" or "autism" (Lord et al., 2000) were characterized as not meeting the criteria for a research-reliable autism diagnosis (i.e., have a research diagnosis that does not confirm their reported community diagnosis).

To align more closely with changes to ASD criteria in the DSM-5, especially with regard to identifying sensory issues, the authors of the ADOS devised a new diagnostic algorithm for Module 4 (Hus & Lord, 2014). This new

algorithm now includes restricted and repetitive behaviors (RRBs) as part of the total score. To determine whether exclusion rates by sex were driven by the absence of RRBs in the total cut-off scores, we rescored all ADOSes for which we had complete item-level data in accordance with the new 2014 algorithm (Hus & Lord, 2014) and re-ran our analyses (total $n = 142$, 93 males, 49 females). New algorithm scores were not able to be calculated for three participants (one female and two males) whose itemized ADOS scores were unavailable. After rescored using the new algorithm, $n = 1$ female and $n = 5$ males who previously met criteria no longer met criteria, and $n = 6$ females and $n = 6$ males who previously did not meet criteria now met criteria.

Statistical analysis

Statistical analyses were conducted with STATA software (Statacorp LLC, 2019). Two-sample t tests were conducted to determine whether ADOS scores differed by sex (two-tailed $p < 0.05$). We also conducted a two-sample t -test to determine whether there were sex differences by age (defined in years at the time at which the ADOS was administered). Chi-squared tests were conducted to assess whether rates of inclusion in research differed by sex. Lastly, logistic regression analyses were used to determine the relationship between ADOS severity and sex (standardizing the ADOS Communication and Reciprocal Social Interaction Total score, and controlling for age).

Comparison databases

We compared sex ratios of participants in our database with participant sex ratios in eight large publically-available datasets commonly used in autism research (Table 1). The process used to confirm autism diagnosis differed across the datasets, with the SPARK and Warrier et al. (2020) databases relying on community diagnosis and the ABIDE I and II databases using ADOS to confirm diagnosis (details below). Datasets that used self-report as the primary measure of autism status often included additional questions to solicit more information about the diagnosis (e.g., "What year was the diagnosis made?", "Who made the diagnosis?", "Name the type of provider", etc.).

SPARK (Simons Foundation Powering Autism Research for Knowledge)

The SPARK database currently consists of almost 100,000 autistic individuals (>15,000 adults) and includes phenotypic as well as genetic data (The SPARK Consortium 2018). Autism diagnosis is ascertained via a survey question inquiring about diagnosis source (e.g., pediatrician,

TABLE 1 Summary of comparison datasets and diagnostic confirmation methods

Comparison dataset	Method to confirm autism status for research	Total ASD <i>N</i>	Total ASD males	Total ASD females
SPARK	SR: Formal community diagnosis + validity + confound flags + exclude for “cognitive impairment”	12,212	7708	4504
Warrier et al., 2020 ^a				
Channel 4 (C4)	SR: “Are you autistic?”	27,251	13,317	13,934
Musical Universe (MU)	SR: Formal diagnosis from a professional	1031	666	365
LifeLines	SR: “Do you have an autism diagnosis?” + year of diagnosis	436	252	184
IMAGE	SR: Formal community diagnosis + additional questions + documentation of diagnosis from professional	330	177	153
Autism Physical Health Survey (APHS)	SR: Formal community diagnosis + additional questions	949	387	562
ABIDE I ^b	ADOS Module 4	123	109	14
ABIDE II ^b	ADOS Module 4	91	80	11
Combined total		42,423	22,696	19,727

Note: Methods to confirm autism diagnostic status for research purposes varied across the comparison datasets. Datasets that used self-report as the primary measure of autism status often included additional questions to solicit more information about the diagnosis (e.g., What year was the diagnosis made? Who made the diagnosis? Name the type of provider, etc.). Several data sets also included neurotypical participant groups, which were not considered in our analyses. For the purposes of the current analyses, “Total *N*” and *N*’s listed for males and females reported in this table refer to the samples sizes of only autistic participants from that dataset. Abbreviations: ABIDE, autism brain imaging data exchange; ADOS, autism diagnostic observation schedule; ASD, autism spectrum disorder; SR, self-report.

^aThe Warrier et al., 2020 sample is comprised of five individual datasets, all of which included information about gender identity (asking participants to self-identify as cis-gender, transgender/gender-diverse). For the purposes of the current analysis and to allow more direct comparisons with other datasets, we only included cis-gender participants from the samples reported in Warrier et al., 2020.

^bIn the ABIDE datasets, sample sizes reported here collapse across sites that used a Module 4 ADOS to determine inclusion, and exclude any sites that explicitly recruited male-only samples.

psychologist, etc.). As part of the available data, SPARK also asks questions meant to provide increased confidence in the diagnosis. SPARK then calculates an “ASD validity flag” that identifies participants for whom diagnostic status might be more uncertain (e.g., diagnosed prior to 15 months of age, never accessed services for autism, had a diagnosis that was rescinded). To match the age range used in our original MIT database analysis, we included all adults age 16 years and above with a reported community diagnosis of autism, autism spectrum disorder, Asperger’s, or Pervasive Developmental Disorder. We excluded participants who had an “age validity flag,” an “ASD validity flag,” or an “ASD confound flag” as determined by SPARK, defined as confounding medical or psychiatric diagnoses (e.g., diagnosis of schizophrenia, other psychosis, or schizoaffective disorder (participant self-report); prematurity (gestational age < 28 weeks); blindness; fetal alcohol syndrome/alcohol or drug exposure in mother’s pregnancy; insufficient oxygen at birth with NICU stay; bleed into the brain; cognitive delays or impairment due to another medical condition or exposure (such as brain injury, stroke, lead poisoning, FAS, HIV, radiation, hydrocephalus, brain tumor, drug effects, etc.); brain infection such as bacterial meningitis or encephalitis). Additionally, we excluded participants with an unspecified diagnostic source (e.g., “not sure” who made the diagnosis) and those with a cognitive impairment at time of

enrollment. This resulted in a total sample size of $n = 12,212$ individuals (7708 males, 4504 females; mean age = 26.09 years, range = 16 to 86 years).

Sample reported by Warrier and colleagues

A manuscript by Warrier et al. (2020), summarizes several datasets originally intended to examine rates of autism in transgender and gender-diverse individuals, collapsing across five separate UK-based publically-available datasets (Channel 4, Musical Universe, LifeLines, IMAGE, Autism Physical Health Survey). Across each of these separate datasets, individuals were asked to self-report their autism diagnosis. The datasets differed in the manner in which they ascertained autism diagnostic status (Table 1). The total sample size of autistic individuals across all datasets included in Warrier et al. was $n = 29,997$ (14,799 males; 15,198 females, see Warrier et al., 2020 for additional demographic details).

Autism brain imaging data exchange I

The autism brain imaging data exchange I (ABIDE I) database consists of 17 independent research sites that have publicly shared imaging and phenotypic data. A

majority of these sites (14 out of 17) utilize the ADOS as part of their inclusion criteria to verify diagnostic status. We calculated sex ratios across ABIDE I to assess whether ADOS usage was associated with sex ratios comparable to those we obtained after ADOS administration in the MIT database. We excluded one site that intentionally recruited male-only samples. From the remaining 13 sites, we included all individuals with a Module-4 ADOS Total Score (item level scores were unavailable) who were 16 years and older. This resulted in a total sample size of $n = 123$ individuals (109 males, 14 females; Mean age 25.26 ± 7.78 ; ranging in age from 16 to 55 years).

Autism brain imaging data exchange II

The autism brain imaging data exchange II (ABIDE II) database consists of 19 independent research sites all of which mention the use of the ADOS as part of their inclusion criteria and confirmatory diagnostic procedure. We excluded two sites that explicitly reported recruiting an all-male sample, and we included only individuals with a Module-4 ADOS Total Score who were 16 years and older. This resulted in a total sample of $n = 91$ (80 males; 11 females; Mean age = 27.71 ± 12.42 ; ranging in age from 16 to 62 years).

RESULTS

Sex differences in assessment scores and exclusion rates for autism research

We first assessed whether sex differences existed among individuals recruited for participation in the MIT Autism Research Participant Database by comparing diagnostic scores on the ADOS Module 4 between males and females. Importantly, all individuals recruited had a community diagnosis of ASD. Females had lower scores on both ADOS Communication ($t[143] = 2.93, p = 0.004$) and Reciprocal Social Interaction ($t[143] = 3.25, p = 0.001$) sub-scales compared to males. There were no sex differences on the ADOS Stereotyped Behaviors and Restricted Interests sub-scale ($t[143] = 0.65, p = 0.514$)

(Table 2). To quantify the extent to which ADOS scores were predictive of sex, we conducted logistic regressions (controlling for age as community diagnosed females were older than community diagnosed males, $t[143] = 2.53, p = 0.012$). Higher ADOS scores were predictive of sex, with each increase of one standard deviation in ADOS score resulting in an almost twofold increase in the probability of being male (odds ratio = 1.80; $p = 0.004$; LCI = 1.20, UCI = 2.69).

These sex differences in ADOS scores impacted the final sex ratio of eligible participants meeting inclusion criteria compared to the recruited sex ratio (Figures 1 and 2), as females were more likely to fall below the diagnostic cut-off scores. After ADOS administration, a greater proportion of females were excluded from further research participation than males (50% of community diagnosed females [$n = 25$] were excluded from further research participation compared with only 19% of community diagnosed males [$n = 18$] [total $n = 145$, $X^2(1) = 15.14, p < 0.001$]). In the final post-ADOS sample ($n = 102$), only 50% of females ($n = 25$) met criteria on the ADOS for “autism” or “autism spectrum” classification, compared with 81% of males ($n = 77$). This shifted the sex ratio (males: females) from 1.9:1 in the recruited sample to 3.1:1 in the post-ADOS sample.

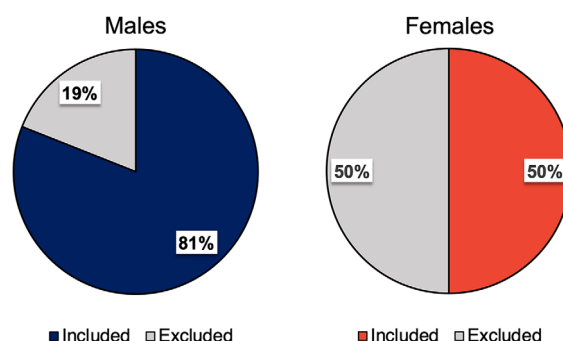


FIGURE 1 Proportion of community-diagnosed adults excluded from research following confirmatory diagnostic assessment. Relative rates of inclusion and exclusion in females versus males recruited for research with a community diagnosis. The “Excluded” percentage indicates the proportion of participants who did not meet cutoff scores for autism or autism spectrum disorder on the autism diagnostic observation schedule (ADOS) and were therefore excluded from research studies. Blue, male; Red, female

TABLE 2 Comparison of ADOS scores by sex

ADOS subscales	Male mean (SD)	Female mean (SD)	<i>t</i> -value	<i>p</i> -value	Hedges' <i>g</i>
Communication (Comm)	3.08 (1.62)	2.28 (1.47)	2.93	0.004	0.51
Reciprocal social interaction (RSI)	6.02 (2.65)	4.48 (2.84)	3.25	0.001	0.56
Comm and RSI	9.11 (3.89)	6.76 (3.84)	3.47	<0.001	0.60
Stereotypical behaviors and restricted interests	1.60 (1.50)	1.42 (1.72)	0.65	0.514	0.11

Note: *p*-values are based on two sample *t* tests between the total sample of males and females ($n = 145$). Hedges' *g* is an effect size measure typically used for unequal sample sizes.

Abbreviations: ADOS, autism diagnostic observation schedule; SD, standard deviation.

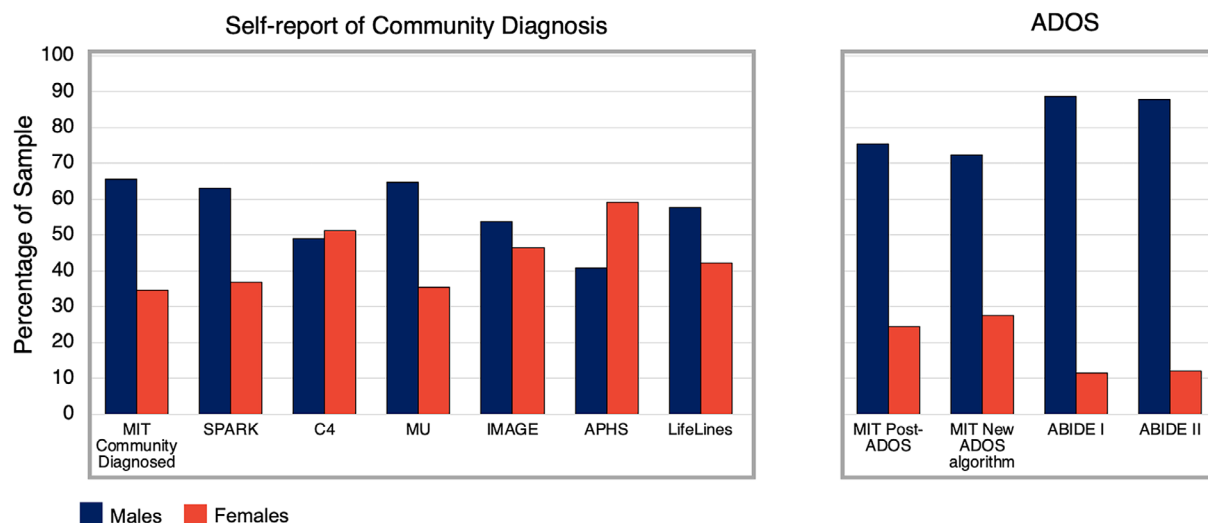


FIGURE 2 Percentage of sample by sex in databases using self-report of community autism diagnosis in comparison to databases that used the autism diagnostic observation schedule (ADOS) to confirm autism diagnosis. The MIT community diagnosed sample ($n = 50$ female; $n = 95$ male) and SPARK database sample ($n = 4504$ female; $n = 7708$ male) consisted of autistic individuals who self-reported a community diagnosis of autism. The MIT post-ADOS sample ($n = 25$ female; $n = 77$ male) consists of participants who self-reported an autism diagnosis which was subsequently confirmed by the ADOS. The ABIDE I dataset ($n = 14$ female; $n = 109$ male) and ABIDE II dataset ($n = 11$ female; $n = 80$ male) consist of participants who self-reported an autism diagnosis and had an ADOS Module 4 total score. The Warrier et al. study consisted of five separate datasets (Channel 4, Musical Universe, LifeLines, IMAGES, Autism Physical Health Survey), all of which used community diagnosis to determine eligibility. Each dataset used slightly different methods for ascertaining autism status (see Table 1 for details). Blue, male; Red, female

To determine whether new algorithms meant to provide better alignment with DSM-5 diagnostic criteria for ASD might result in greater inclusion of females, we calculated exclusion rates by sex using the revised ADOS-2 algorithms (Hus & Lord, 2014). These algorithms include sensory symptoms and RRBs as part of the total score (Hus & Lord, 2014). The new algorithm resulted in a net percentage increase of both females and males meeting criteria on the ADOS ($n = 29/49$ females included [59%], $n = 76/93$ males included [82%]), as well as a moderate decrease in the final male-to-female ratio (2.6:1 compared with 3.1:1 using the original algorithm). Exclusion rates decreased with the new algorithm for both males and females, revealing a greater decrease for females (18% males excluded vs. 19% when using the original algorithm; 41% females excluded as compared with 50% when using the original algorithm). Interestingly, rescoring with the new algorithm also resulted in changes in diagnostic status in a few individuals (i.e., participants who previously met criteria no longer met, $n = 1$ female, $n = 5$ males). However, even when using the revised algorithm, there were significant differences in exclusion rates by sex (total $n = 142$, $X^2[1] = 8.46$, $p = 0.004$), with a greater proportion of females being excluded post-ADOS than males (Figure 2).

Comparison to large, publically-available autism datasets

We next examined whether our findings paralleled sex ratios in large, publically-available autism datasets. We

identified datasets that either used (a) community diagnosis (reported by the participant or family member of the participant) to characterize participants (The SPARK Consortium; five separate datasets as reported in Warrier et al., 2020) or (b) relied at least in part on the ADOS to determine inclusion (ABIDE I; ABIDE II) (see Table 1). Sex ratios were calculated for each of these datasets. The proportion of females with an ASD designation was markedly smaller in datasets wherein the ADOS was used to determine research eligibility (male:female ratios, ABIDE I, 7.8:1; ABIDE II, 7.3:1) compared with datasets that employed self-reported community diagnosis (SPARK, 1.7:1; C4, 0.95:1; MU, 1.8:1; IMAGE, 1.1:1; LifeLines, 1.3:1; APHS, 0.68:1) (Figure 2). Studies using samples based on community diagnosis often included additional questions about that diagnosis, but the specific questions varied across samples and it is thus unclear what role these questions had for inclusion/exclusion. Our findings indicate that the sex ratio of the MIT community diagnosed sample closely parallels the sex ratio of datasets that used self-reported community diagnosis to determine inclusion. In contrast, the sex ratio of the MIT post-ADOS sample was similar to that of the ABIDE samples, which also utilized the ADOS to confirm diagnosis of participants.

DISCUSSION

The current study is the first to empirically assess how differing research practices for confirming diagnosis and determining inclusion ultimately affect the number of

ASD females deemed eligible to participate in research. Across our sample and several other large datasets, using community diagnosis to determine eligibility resulted in more equal sex ratios between females and males. The use of common diagnostic assessment measures such as the ADOS had a disproportionate effect on the exclusion of autistic females in research relative to autistic males. We suggest that these procedures may contribute to the underrepresentation of females in autism research, and provide evidence for a “leaky” recruitment-to-research pipeline for autistic females.

Confirmatory diagnostic assessments affect the male-to-female sex ratio in ASD research

In the current study, the MIT community-diagnosed sample showed sex ratios of ~2:1 (male:female) prior to implementation of confirmatory diagnostic assessments. ADOS usage to confirm diagnosis resulted in a disproportionate exclusion of females as compared to males. This was driven by lower social-communication ADOS scores in females, several of which fell below cut-off scores for an autism designation. Sex differences in ADOS scores and exclusion rates persisted even when using updated scoring algorithms designed to reflect the most current understanding of ASD criteria in the DSM-5, specifically the inclusion of RRBs in the ADOS total score (Hus & Lord, 2014). Similar results emerged across eight large publically-available datasets that used different measures to determine inclusion (community diagnosis vs. ADOS), with some datasets that relied on community diagnosis even showing a greater number of females than males. Importantly, although ratios of 2:1 (male:female) in our original recruited sample may seem anomalous given the existing literature and consensus estimates commonly reporting sex ratios of around 4:1, we replicated these ratios in much larger international and publically available datasets (e.g., SPARK, C4, Life-Line, IMAGE, MU, and APhS) all of which relied on community diagnosis. Our findings highlight that relying on cut-off scores on confirmatory diagnostic assessments without expert clinical consensus are one potential contributor to the small sample sizes of autistic females in research, and also potentially contribute to the variable sex ratios found in the ASD literature (Loomes et al., 2017).

These results are consistent with several studies from the past decade suggesting that gold-standard instruments such as the ADOS poorly identify autistic females (Lai et al., 2011; Ratto et al., 2018; Rynkiewicz & Łucka, 2018; Tillmann et al., 2018). Why might these tools have these effects? Of most import, several ASD trait and diagnostic measures, including the ADOS, were standardized predominantly on males and do not offer sex-specific norms (Baron-Cohen et al., 2001; Lord et al., 2000). Studies examining the validity of the ADOS,

and several other commonly used diagnostic assessments, use primarily male samples (Medda et al., 2019) and in some cases, exclusively male samples. Therefore, females are more likely to be classified as autistic by the ADOS when they display behaviors more similar to their male peers, despite suggestions that the female ASD phenotype may be distinct from that of males (Medda et al., 2019). In addition, autistic females are more likely than their male counterparts to mask autistic traits in social and clinical settings, also known as camouflaging (Lai et al., 2017; Tubío-Fungueiriño et al., 2021). Further, females' self-report of their own autistic traits is much higher than when measured by observational measures such as the ADOS (Cook et al., 2021; Hull et al., 2020; Lai et al., 2017). Tools such as the ADOS are not designed to detect camouflaging, resulting in lower scores (fewer challenges detected) for females, potentially exacerbating sex-based exclusions. These tools may also be vulnerable to gender-based biases on the part of the clinician or administrator. For example, social communication skills are perceived as being better in autistic girls than boys (despite equivalent symptom severity in this domain), resulting in better first impressions (Cola et al., 2022). One study reported that autistic females used more social words than males during clinical assessments, even when matched for symptom severity. Crucially, individuals who used more social words received lower ADOS scores by clinicians (Cola et al., 2022). These diagnostic tools are also less accurate in capturing certain subgroups or phenotypes, such as autistic females with high IQs (Ratto et al., 2018). Lastly, previous reports have found that repetitive behaviors and restricted interests (RRBs) may be more predictive of autism diagnostic status than social communication scores (Troyb et al., 2016). However, in most current research practice, only social communication scores are used to determine cut-offs (although see Hus & Lord, 2014, as of 2014 the new ADOS algorithm includes RRBs in the total score). Here, we report robust sex differences in social communication symptoms, but no sex differences in RRBs as measured by the ADOS. Concomitantly, use of the new algorithm (which incorporates RRBs into the total cut-off score) did slightly increase the number of females meeting criteria for ASD, suggesting that including RRBs may contribute to decreasing the sex difference in ADOS scores driven largely by differences in social communication (but see Lai & Szatmari, 2020 for evidence of sex differences in RRBs). Despite this slight increase, however, use of the new algorithm did not significantly alter inclusion rates by sex.

Apart from the particularly disproportionate effect on females, our data provide evidence that the use of confirmatory diagnostic assessments narrows the research sample overall, excluding large proportions of *both* females and males with a community autism diagnosis (e.g., 19% of community diagnosed males in our recruited sample were excluded based on ADOS cutoffs). This suggests

that the ADOS, whether used as an assessment or diagnostic confirmation tool, might only be capturing a certain part of the autistic population, which could be distinct from the broader community diagnosed sample. While some argue that a more homogenous autism sample might actually strengthen research findings (Mottron, 2021), a more homogeneous sample may not adequately capture characteristics representative of the full autism spectrum. It is unclear how accessing a certain proportion of the autism population affects research findings, and how results arising from this research generalize to autism in general.

Underepresentation of females in autism research begins before recruitment: Barriers to obtaining diagnosis

There are several factors that may limit participation of autistic females in research that occur *prior* to the research process. Females face greater barriers to obtaining an autism diagnosis which may bias the sex ratio and result in a smaller pool of females before the point of research (for a review, see Lockwood Estrin et al., 2020). These barriers include differing phenotypic presentation that may not align with conceptualizations of autism (e.g., RRBs that more closely match societal norms), perception of autism as a male disorder, and social norms surrounding social communication abilities in females (Cola et al., 2022; Hiller et al., 2014). For example, diagnosticians report that they perceive diagnostic assessments of ASD to be more challenging when the client is female, and attribute this difficulty to their perceptions of incongruence between current ASD tools (and conceptualization) and female presentation (Tsirgiotis et al., 2021). Indeed, one study found that despite no difference in the duration of diagnostic assessments, females were still less likely to receive a diagnosis and parents of ASD girls had to exaggerate symptoms in order to get a diagnosis for their daughters (Rutherford et al., 2016).

These issues may be influenced by development. Although not explicitly explored in this study, male:female ratios in community-based samples changed markedly as a function of age (see Data S1 for statistics, table, and a brief discussion). These age-related discrepancies may speak to biases in diagnostic practices and gendered social norms. For instance, due to societal norms, repetitive or stereotyped play patterns in girls may, on the surface, seem more appropriate than those in boys (focus on repetitive play with dolls vs. wheels on a truck; [Giarelli et al., 2010; Hiller et al., 2016]). This, in turn, may lead to differences in perception, identification, and scoring of RRBs in standardized assessments, and ultimately later age of diagnosis in females—a commonly observed occurrence (Harrop et al., 2021). Indeed, girls are less likely to meet diagnostic thresholds than boys, despite having equally high levels of autistic traits

(Dworzynski et al., 2012; Kalb et al., 2022; Mo et al., 2021), and teachers report significantly fewer concerns about social skills in girls than boys (Hiller et al., 2014). This suggests that females may not be referred for diagnosis as often, or as early, as males. Consistent with this, we found that community-diagnosed females in our sample tended to be older on average than males (Table S1, Figure S1).

Lastly, the numbers of autistic females in research may be constrained by true prevalence of autism in females. However, given issues raised in this article and elsewhere, the true prevalence is difficult to determine. These diagnostic and conceptual difficulties, even prior to research, suggest that additional sex-based norming and tool evaluation is needed to improve diagnosis in autistic females.

Implications and considerations for future research

The potential underrepresentation of autistic females in autism research has multiple implications for the study, characterization, and acceptance of autistic women in society, as well as access to services. Scientifically, consistently small sample sizes of autistic females make it difficult to fully understand autism in females. Underrepresentation in research not only perpetuates the perception that autism is a male disorder, but may create a cycle in which basic science questions and investigations of autism are constrained to specific phenotypes based on what has already been studied in the previous literature (Figure 3). Our research suggests that females who are ultimately included in research (i.e., whose diagnoses are confirmed using the ADOS) exhibit a specific phenotype that may or may not represent autism in females more broadly. Crucially, exclusion begets more exclusion. Because autism tends to be viewed as a characteristically male disorder, researchers are often excused from including adequate samples of females in research. As a striking example, in a sample of over 1400 studies focusing on the brain structure and function in autism, only 4 studies focused on female-only samples compared to 434 using male-only samples (Mo et al., 2021). This may, in turn, inform the topic and focus of future autism research, constrain recruitment efforts, and even impede the development of diagnostic tools. Lastly, an important goal of research is to ultimately inform the delivery of services and support to individuals with ASD. Indeed, non-male identifying autistic individuals (both females and individuals with other non-male gender identities) report greater difficulty accessing services than males (Koffer Miller et al., 2022), and autistic females tend to feel ignored or unseen (Bargiela et al., 2016).

How might research adapt to address these issues? To begin, community diagnosis reported by the participant or family member, along with confirmatory details and

Interactions between Research, Diagnostic, and Recruitment Practices Magnify the “Leaky Pipeline” Effect

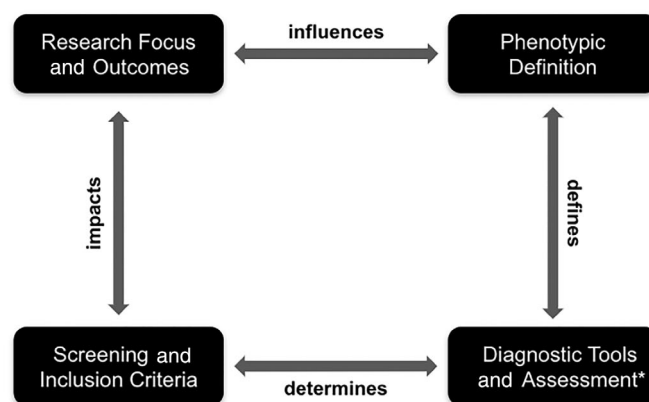


FIGURE 3 Diagram of the interactions between research, diagnostic, and recruitment practices. If females are excluded from any part of these processes it magnifies the degree of discrepant exclusion rates. The outcomes of research may directly or indirectly contribute to biased perceptions and diagnostic practices in autism (and vice versa). Reduced representation of females in autism research (due to focus on males, recruitment of primarily male samples, etc.) may increase the general perception that autism is primarily a male disorder, and strengthens the idea that the ASD male phenotype is the phenotypic template on which diagnostic definitions should be based. These perceptions have knock-on effects on the construction of diagnostic tools and assessments: Because these tools are normed in primarily male samples, they act to further entrench biased perceptions about what may or may not reflect true autistic behavior. These issues have basic science and translational implications: Our understanding of autism is unlikely to be entirely accurate without adequate representation of females, and fewer females in research lessen the probability that any results will generalize broadly to autism or improve outcomes for autistic individuals. *Refers to the development of assessment tools, and their implementation in diagnosis and determination of research eligibility.

procedures, can serve as a valuable consideration for researchers to increase representation of females in research. Crucially, researchers do not have to forgo *all* diagnostic confirmation when using self-report of community diagnosis. Several large autism databases ask participants to provide further information about their diagnosis (e.g., copy of the diagnostic report, source of diagnosis). Further, there may be cases in which additional confirmatory procedures are beneficial (for instance when an extended period of time has elapsed since participant's last evaluation, in light of research showing that a subset of autistic children may “outgrow” their diagnoses, [Anderson et al., 2014; Fein et al., 2013]). In addition, researchers can report pre- and post-confirmatory diagnostic assessment sample sizes (how many individuals were originally recruited, how many were excluded as a result of diagnostic confirmation, which criteria are particularly exclusionary, whether there was discrepancy between community diagnosis and research-based diagnostic confirmation measures, and how these numbers differ by sex and gender). This reporting, coupled with a critical examination of exclusion criteria, will inform researchers about exclusion criteria that are sex-related. Exclusion criteria may interact with sex and gender identity. Given higher rates of gender diversity in ASD than in non-ASD individuals, it is especially important to assess the effects of diagnostic confirmation by both gender and sex (see Warriar et al., 2020 for discussion of gender identity in autistic individuals, Beltz et al., 2019 for recommendations on best practices for

conducting research in sex differences, and Strang et al., 2020 on the importance of considering both sex and gender in autism research). For researchers using the ADOS or other confirmatory assessments, it is important to understand how interactions between examiner biases and sex differences can play a role in the identification of autism traits that may qualitatively differ in females and are closer to societally accepted behavioral norms. Researchers can also consider using *sex-independent* tasks and trait assessments that differ between autistic and non-autistic populations, but do not show sex differences (e.g., the Reading the Mind in the Eyes Task shows no sex differences in ASD, Baron-Cohen et al., 2015). Additionally, measurements should be used in concert with clinical judgment and self-report. Our research and that of others suggests that strict categorical cut-offs exclude females to a greater degree, and confirmatory diagnostic tools are less accurate in identifying individuals near cut-off borders (Charman & Gotham, 2013). Lastly, journals and autism research societies may also play a role by discouraging reviewers from penalizing manuscripts that rely on community diagnosis. Together, these considerations could contribute to increasing representation of females in autism research.

Limitations and conclusions

The current results should be interpreted within the context of a few limitations. For instance, although all

individuals had a community diagnosis of ASD, we did not ascertain the source of, or age of, community diagnosis in all participants. Second, inclusion was determined using the ADOS and we cannot, therefore, speak to the applicability of other commonly used confirmatory diagnostic measures (e.g., the ADI). In addition, the current study only examined adolescents or adults who were verbal (administered the Module 4 of the ADOS) and it is unclear whether similar results would hold in samples that are non-speaking, minimally verbal, or cognitively impaired. Further, the included comparison databases differed in the extent to which co-occurring medical and psychiatric issues were characterized, as well as how strictly ADOS cut-offs were applied or reconciled with clinical judgment. Moreover, here, we used binary sex definitions, which may not capture gender-based differences, including those specific to nonbinary-identifying and transgender individuals. Lastly, it is possible that basic science research relies more on strict ADOS cut-offs than does clinically-oriented research or research teams in which a clinician is present, and it is unclear what proportion of autism research studies adhere to rigid cut-offs or use the ADOS to confirm diagnosis.

Despite these limitations, we find robust evidence that confirmatory diagnostic assessments commonly used in autism research may contribute to the small sample sizes of females in autism research. By examining both our sample and more than 42,000 autistic individuals from eight comparison datasets, we find that utilizing self-report of community diagnosis can contribute to dramatically lower sex ratios in autism research. Our analyses reveal that even datasets that explicitly obtain diagnostic reports to confirm autism status had more balanced sex ratios than those that used only the ADOS. Strong reliance on such measures may play a role in perpetuating the disproportionate exclusion of females in autism research. Across females and males, future research should characterize both the brain and behavior in autistic individuals with community diagnoses, comparing those who are and who are not excluded by standardized measures. These considerations could play an important role in increasing representation of females in research (see Figure 3).

ACKNOWLEDGMENTS

The authors thank the individuals who participated in autism research at MIT and acknowledge support from the Hock E. Tan and K. Lisa Yang Center for Autism Research and the Simons Center for the Social Brain at MIT for the creation of the MIT Autism Research Participant Database, support of autism recruitment at MIT, and support of the autism research coordinator (Cindy E. Li). The authors would further like to acknowledge funding support from the Simons Center for the Social Brain at MIT (postdoctoral fellowship to Anila M. D'Mello and Targeted Project Grant: Predictive Processes in Autistic and Neurotypical Individuals: a

behavioral, neural and developmental investigation to John D.E. Gabrieli) and National Institutes of Mental Health (F32 MH117933 to Anila M. D'Mello). The authors also thank Drs. Pawan Sinha and Robert Joseph for comments on the manuscript.

DATA AVAILABILITY STATEMENT

De-identified data from the Autism Research Participant Database at MIT are available upon request. Data from comparison databases are publicly available, or available via request (e.g., SPARK database access requires application).

ETHICS STATEMENT

All procedures were approved by the MIT Institutional Review Board (Committee on the Use of Humans as Experimental Subjects).

ORCID

Anila M. D'Mello  <https://orcid.org/0000-0001-6653-7802>

Isabelle R. Frosch  <https://orcid.org/0000-0001-9508-2747>

REFERENCES

- Anderson, D. K., Liang, J. W., & Lord, C. (2014). Predicting young adult outcome among more and less cognitively able individuals with autism spectrum disorders. *Journal of Child Psychology and Psychiatry*, 55(5), 485–494. <https://doi.org/10.1111/jcpp.12178>
- Bargiela, S., Steward, R., & Mandy, W. (2016). The experiences of late-diagnosed women with autism spectrum conditions: An investigation of the female autism phenotype. *Journal of Autism and Developmental Disorders*, 46(10), 3281–3294. <https://doi.org/10.1007/s10803-016-2872-8>
- Barnard-Brak, L., Richman, D., & Almekdash, M. H. (2019). How many girls are we missing in ASD? An examination from a clinic- and community-based sample. *Advances in Autism*, 5(3), 214–224. <https://doi.org/10.1108/AIA-11-2018-0048>
- Baron-Cohen, S., Bowen, D. C., Holt, R. J., Allison, C., Auyeung, B., Lombardo, M. V., Smith, P., & Lai, M.-C. (2015). The “reading the mind in the eyes” test: Complete absence of typical sex difference in ~400 men and women with autism. *PLoS One*, 10(8), e0136521. <https://doi.org/10.1371/journal.pone.0136521>
- Baron-Cohen, S., Wheelwright, S., Skinner, R., Martin, J., & Clubley, E. (2001). The autism-spectrum quotient (AQ): Evidence from Asperger syndrome/high-functioning autism, males and females, scientists and mathematicians. *Journal of Autism and Developmental Disorders*, 31(1), 5–17. <https://doi.org/10.1023/A:1005653411471>
- Bastiaansen, J. A., Meffert, H., Hein, S., Huizinga, P., Ketelaars, C., Pijnenborg, M., Bartels, A., Minderaa, R., Keyers, C., & de Bildt, A. (2011). Diagnosing autism spectrum disorders in adults: The use of autism diagnostic observation schedule (ADOS) module 4. *Journal of Autism and Developmental Disorders*, 41(9), 1256–1266. <https://doi.org/10.1007/s10803-010-1157-x>
- Beltz, A. M., Beery, A. K., & Becker, J. B. (2019). Analysis of sex differences in pre-clinical and clinical data sets. *Neuropsychopharmacology*, 44(13), 2155–2158. <https://doi.org/10.1038/s41386-019-0524-3>
- Bottema-Beutel, K., Kapp, S. K., Lester, J. N., Sasson, N. J., & Hand, B. N. (2021). Avoiding Ableist language: Suggestions for autism researchers. *Autism in Adulthood*, 3(1), 18–29. <https://doi.org/10.1089/aut.2020.0014>

- Burrows, C. A., Grzadzinski, R. L., Donovan, K., Stallworthy, I. C., Rutsohn, J., St. John, T., Marrus, N., Parish-Morris, J., MacIntyre, L., Hampton, J., Pandey, J., Shen, M. D., Botteron, K. N., Estes, A. M., Dager, S. R., Hazlett, H. C., Pruett, J. R., Schultz, R. T., Zwaigenbaum, L., ... Elison, J. T. (2022). A data driven approach in an unbiased sample reveals equivalent sex ratio of autism spectrum disorder associated impairment in early childhood. *Biological Psychiatry*. <https://doi.org/10.1016/j.biopsych.2022.05.027>
- Charman, T., & Gotham, K. (2013). Measurement issues: Screening and diagnostic instruments for autism spectrum disorders - lessons from research and practise. *Child and Adolescent Mental Health*, 18(1), 52–63. <https://doi.org/10.1111/j.1475-3588.2012.00664.x>
- Cola, M., Yankowitz, L. D., Tena, K., Russell, A., Bateman, L., Knox, A., Plate, S., Cubit, L. S., Zampella, C. J., Pandey, J., Schultz, R. T., & Parish-Morris, J. (2022). Friend matters: Sex differences in social language during autism diagnostic interviews. *Molecular Autism*, 13(1), 5. <https://doi.org/10.1186/s13229-021-00483-1>
- Cook, J., Hull, L., Crane, L., & Mandy, W. (2021). Camouflaging in autism: A systematic review. *Clinical Psychology Review*, 89, 102080. <https://doi.org/10.1016/j.cpr.2021.102080>
- Cumin, J., Pelaez, S., & Mottron, L. (2021). Positive and differential diagnosis of autism in verbal women of typical intelligence: A Delphi study. *Autism*, 26, 1153–1164. <https://doi.org/10.1177/13623613211042719>
- Dworzynski, K., Ronald, A., Bolton, P., & Happé, F. (2012). How different are girls and boys above and below the diagnostic threshold for autism spectrum disorders? *Journal of the American Academy of Child & Adolescent Psychiatry*, 51(8), 788–797. <https://doi.org/10.1016/j.jaac.2012.05.018>
- Fein, D., Barton, M., Eigsti, I.-M., Kelley, E., Naigles, L., Schultz, R. T., Stevens, M., Helt, M., Orinstein, A., Rosenthal, M., Troyb, E., & Tyson, K. (2013). Optimal outcome in individuals with a history of autism. *Journal of Child Psychology and Psychiatry*, 54(2), 195–205. <https://doi.org/10.1111/jcpp.12037>
- Fusar-Poli, L., Brondino, N., Politi, P., & Aguglia, E. (2020). Missed diagnoses and misdiagnoses of adults with autism spectrum disorder. *European Archives of Psychiatry and Clinical Neuroscience*, 272, 187–198. <https://doi.org/10.1007/s00406-020-01189-w>
- Giarelli, E., Wiggins, L. D., Rice, C. E., Levy, S. E., Kirby, R. S., Pinto-Martin, J., & Mandell, D. (2010). Sex differences in the evaluation and diagnosis of autism spectrum disorders among children. *Disability and Health Journal*, 3(2), 107–116. <https://doi.org/10.1016/j.dhjo.2009.07.001>
- Harrop, C., Libsack, E., Bernier, R., Dapretto, M., Jack, A., McPartland, J. C., Van Horn, J. D., Webb, S. J., Pelphrey, K., & Consortium, the G. (2021). Do biological sex and early developmental milestones predict the age of first concerns and eventual diagnosis in autism spectrum disorder? *Autism Research*, 14(1), 156–168. <https://doi.org/10.1002/aur.2446>
- Hiller, R. M., Young, R. L., & Weber, N. (2014). Sex differences in autism spectrum disorder based on DSM-5 criteria: Evidence from clinician and teacher reporting. *Journal of Abnormal Child Psychology*, 42(8), 1381–1393. <https://doi.org/10.1007/s10802-014-9881-x>
- Hiller, R. M., Young, R. L., & Weber, N. (2016). Sex differences in pre-diagnosis concerns for children later diagnosed with autism spectrum disorder. *Autism*, 20(1), 75–84. <https://doi.org/10.1177/1362361314568899>
- Hinshaw, S. P., Nguyen, P. T., O'Grady, S. M., & Rosenthal, E. A. (2022). Annual research review: Attention-deficit/hyperactivity disorder in girls and women: Underrepresentation, longitudinal processes, and key directions. *Journal of Child Psychology and Psychiatry*, 63, 484–496. <https://doi.org/10.1111/jcpp.13480>
- Hull, L., Petrides, K. V., & Mandy, W. (2020). The female autism phenotype and camouflaging: A narrative review. *Review Journal of Autism and Developmental Disorders*, 7(4), 306–317. <https://doi.org/10.1007/s40489-020-00197-9>
- Hus, V., & Lord, C. (2014). The autism diagnostic observation schedule, module 4: Revised algorithm and standardized severity scores. *Journal of Autism and Developmental Disorders*, 44(8), 1996–2012. <https://doi.org/10.1007/s10803-014-2080-3>
- Jack, A., Sullivan, C. A. W., Aylward, E., Bookheimer, S. Y., Dapretto, M., Gaab, N., Van Horn, J. D., Eilbott, J., Jacokes, Z., Torgerson, C. M., Bernier, R. A., Geschwind, D. H., McPartland, J. C., Nelson, C. A., Webb, S. J., Pelphrey, K. A., Gupta, A. R., the GENDAAR Consortium, Bernier, R. A., ... Ventola, P. (2021). A neurogenetic analysis of female autism. *Brain*, 144(6), 1911–1926. <https://doi.org/10.1093/brain/awab064>
- Kalb, L. G., Singh, V., Hong, J. S., Holingue, C., Ludwig, N. N., Pfeiffer, D., Reetzke, R., Gross, A. L., & Landa, R. (2022). Analysis of race and sex bias in the autism diagnostic observation schedule (ADOS-2). *JAMA Network Open*, 5(4), e229498. <https://doi.org/10.1001/jamanetworkopen.2022.9498>
- Koffler Miller, K. H., Cooper, D. S., Song, W., & Shea, L. L. (2022). Self-reported service needs and barriers reported by autistic adults: Differences by gender identity. *Research in Autism Spectrum Disorders*, 92, 101916. <https://doi.org/10.1016/j.rasd.2022.101916>
- Lai, M.-C., Baron-Cohen, S., & Buxbaum, J. D. (2015). Understanding autism in the light of sex/gender. *Molecular Autism*, 6(1), 24. <https://doi.org/10.1186/s13229-015-0021-4>
- Lai, M.-C., Lombardo, M. V., Auyeung, B., Chakrabarti, B., & Baron-Cohen, S. (2015). Sex/gender differences and autism: Setting the scene for future research. *Journal of the American Academy of Child & Adolescent Psychiatry*, 54(1), 11–24. <https://doi.org/10.1016/j.jaac.2014.10.003>
- Lai, M.-C., Lombardo, M. V., Pasco, G., Ruigrok, A. N. V., Wheelwright, S. J., Sadek, S. A., Chakrabarti, B., MRC AIMS Consortium, & Baron-Cohen, S. (2011). A behavioral comparison of male and female adults with high functioning autism Spectrum conditions. *PLoS One*, 6(6), e20835. <https://doi.org/10.1371/journal.pone.0020835>
- Lai, M.-C., Lombardo, M. V., Ruigrok, A. N., Chakrabarti, B., Auyeung, B., Szatmari, P., Happé, F., Baron-Cohen, S., & MRC AIMS Consortium. (2017). Quantifying and exploring camouflaging in men and women with autism. *Autism*, 21(6), 690–702. <https://doi.org/10.1177/1362361316671012>
- Lai, M.-C., & Szatmari, P. (2020). Sex and gender impacts on the behavioural presentation and recognition of autism. *Current Opinion in Psychiatry*, 33(2), 117–123. <https://doi.org/10.1097/YCO.0000000000000575>
- Lockwood Estrin, G., Milner, V., Spain, D., Happé, F., & Colvert, E. (2020). Barriers to autism spectrum disorder diagnosis for young women and girls: A systematic review. *Review Journal of Autism and Developmental Disorders*, 8, 454–470. <https://doi.org/10.1007/s40489-020-00225-8>
- Loomes, R., Hull, L., & Mandy, W. P. L. (2017). What is the male-to-female ratio in autism spectrum disorder? A systematic review and meta-analysis. *Journal of the American Academy of Child & Adolescent Psychiatry*, 56(6), 466–474. <https://doi.org/10.1016/j.jaac.2017.03.013>
- Lord, C. (2012). A multisite study of the clinical diagnosis of different autism spectrum disorders. *Archives of General Psychiatry*, 69(3), 306–313. <https://doi.org/10.1001/archgenpsychiatry.2011.148>
- Lord, C., Risi, S., Lambrecht, L., Cook, E. H., Leventhal, B. L., DiLavore, P. C., Pickles, A., & Rutter, M. (2000). The autism diagnostic observation schedule-generic: A standard measure of social and communication deficits associated with the spectrum of autism. *Journal of Autism and Developmental Disorders*, 30, 205–223.
- Mandy, W., Chilvers, R., Chowdhury, U., Salter, G., Seigal, A., & Skuse, D. (2012). Sex differences in autism spectrum disorder: Evidence from a large sample of children and adolescents. *Journal of*

- Autism and Developmental Disorders*, 42(7), 1304–1313. <https://doi.org/10.1007/s10803-011-1356-0>
- Medda, J. E., Cholemkery, H., & Freitag, C. M. (2019). Sensitivity and specificity of the ADOS-2 algorithm in a large German sample. *Journal of Autism and Developmental Disorders*, 49(2), 750–761. <https://doi.org/10.1007/s10803-018-3750-3>
- Mo, K., Sadoway, T., Bonato, S., Ameis, S. H., Anagnostou, E., Lerch, J. P., Taylor, M. J., & Lai, M.-C. (2021). Sex/gender differences in the human autistic brains: A systematic review of 20 years of neuroimaging research. *NeuroImage: Clinical*, 32, 102811. <https://doi.org/10.1016/j.nicl.2021.102811>
- Mottron, L. (2021). A radical change in our autism research strategy is needed: Back to prototypes. *Autism Research*, 14(10), 2213–2220. <https://doi.org/10.1002/aur.2494>
- Posserud, M., Skretting Solberg, B., Engeland, A., Haavik, J., & Klungsoyr, K. (2021). Male to female ratios in autism spectrum disorders by age, intellectual disability and attention-deficit/hyperactivity disorder. *Acta Psychiatrica Scandinavica*, 144, 635–646. <https://doi.org/10.1111/acps.13368>
- Ratto, A. B., Kenworthy, L., Yerys, B. E., Bascom, J., Wieckowski, A. T., White, S. W., Wallace, G. L., Pugliese, C., Schultz, R. T., Ollendick, T. H., Scarpa, A., Seese, S., Register-Brown, K., Martin, A., & Anthony, L. G. (2018). What about the girls? Sex-based differences in autistic traits and adaptive skills. *Journal of Autism and Developmental Disorders*, 48(5), 1698–1711. <https://doi.org/10.1007/s10803-017-3413-9>
- Rea, H. M., Øien, R. A., Shic, F., Webb, S. J., & Ratto, A. B. (2022). Sex differences on the ADOS-2. *Journal of Autism and Developmental Disorders*. <https://doi.org/10.1007/s10803-022-05566-3>
- Rutherford, M., McKenzie, K., Johnson, T., Catchpole, C., O'Hare, A., McClure, I., Forsyth, K., McCartney, D., & Murray, A. (2016). Gender ratio in a clinical population sample, age of diagnosis and duration of assessment in children and adults with autism spectrum disorder. *Autism*, 20(5), 628–634. <https://doi.org/10.1177/1362361315617879>
- Rynkiewicz, A., & Łucka, I. (2018). Autism spectrum disorder (ASD) in girls. Co-occurring psychopathology. Sex differences in clinical manifestation. *Psychiatria Polska*, 52(4), 629–639. <https://doi.org/10.12740/PP/OnlineFirst/58837>
- Shansky, R. M., & Murphy, A. Z. (2021). Considering sex as a biological variable will require a global shift in science culture. *Nature Neuroscience*, 24(4), 457–464. <https://doi.org/10.1038/s41593-021-00806-8>
- Strang, J. F., van der Miesen, A. I., Caplan, R., Hughes, C., daVanport, S., & Lai, M.-C. (2020). Both sex- and gender-related factors should be considered in autism research and clinical practice. *Autism*, 24(3), 539–543. <https://doi.org/10.1177/1362361320913192>
- Tillmann, J., Ashwood, K., Absoud, M., Bölte, S., Bonnet-Brilhault, F., Buitelaar, J. K., Calderoni, S., Calvo, R., Canal-Beitia, R., Canitano, R., De Bildt, A., Gomot, M., Hoekstra, P. J., Kaale, A., McConachie, H., Murphy, D. G., Narzisi, A., Oosterling, I., Pejovic-Milovancevic, M., ... Charman, T. (2018). Evaluating sex and age differences in ADI-R and ADOS scores in a large European multi-site sample of individuals with autism spectrum disorder. *Journal of Autism and Developmental Disorders*, 48(7), 2490–2505. <https://doi.org/10.1007/s10803-018-3510-4>
- Tsirgiotis, J. M., Young, R. L., & Weber, N. (2021). A mixed-methods investigation of diagnostician sex/gender-bias and challenges in assessing females for autism spectrum disorder. *Journal of Autism and Developmental Disorders*. <https://doi.org/10.1007/s10803-021-05300-5>
- Troyb, E., Knoch, K., Herlihy, L., Stevens, M. C., Chen, C.-M., Barton, M., Treadwell, K., & Fein, D. (2016). Restricted and repetitive behaviors as predictors of outcome in autism spectrum disorders. *Journal of Autism and Developmental Disorders*, 46(4), 1282–1296. <https://doi.org/10.1007/s10803-015-2668-2>
- Tubío-Fungueiriño, M., Cruz, S., Sampaio, A., Carracedo, A., & Fernández-Prieto, M. (2021). Social camouflaging in females with autism spectrum disorder: A systematic review. *Journal of Autism and Developmental Disorders*, 51(7), 2190–2199. <https://doi.org/10.1007/s10803-020-04695-x>
- Warrier, V., Greenberg, D. M., Weir, E., Buckingham, C., Smith, P., Lai, M.-C., Allison, C., & Baron-Cohen, S. (2020). Elevated rates of autism, other neurodevelopmental and psychiatric diagnoses, and autistic traits in transgender and gender-diverse individuals. *Nature Communications*, 11(1), 3959. <https://doi.org/10.1038/s41467-020-17794-1>
- Watkins, E. E., Zimmermann, Z. J., & Poling, A. (2014). The gender of participants in published research involving people with autism spectrum disorders. *Research in Autism Spectrum Disorders*, 8(2), 143–146. <https://doi.org/10.1016/j.rasd.2013.10.010>
- World Health Organization (Ed.). (1992). *The ICD-10 classification of mental and behavioural disorders: Clinical descriptions and diagnostic guidelines*. World Health Organization.

SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.

How to cite this article: D'Mello, A. M., Frosch, I. R., Li, C. E., Cardinaux, A. L., & Gabrieli, J. D. E. (2022). Exclusion of females in autism research: Empirical evidence for a “leaky” recruitment-to-research pipeline. *Autism Research*, 15(10), 1929–1940. <https://doi.org/10.1002/aur.2795>