Article: cover letter at end of document.

Estimating the Male to Female Ratio in Autistic Spectrum Disorder, Defining and Calculating the Biases and Finding the True Number of Girls.

Robert McCrossin MB BS MPH FRACP

Paediatrician: Cooroy Family Practice and Beerwah Family Clinic bobmccrossin@gmail.com

Abstract

The ratio of males to females with ASD is generally quoted as 4:1 though it is believed that there are biases preventing females being diagnosed and that the true ratio is lower. These biases have not been clearly identified or quantified. Starting with a clinical dataset of 1711 children <18 years old four different methods were employed in an inductive study to identify and quantify the biases and calculate the proportion of females missed. A mathematical model was constructed to compare the findings with current published data. The true male to female ratio appears to be 3:4. Eighty per cent of females remain undiagnosed at age 18 which has serious consequences for the mental health of young women.

Keywords: autistic spectrum disorder, male to female ratio, biases, young women.

INTRODUCTION

1 Problem statement.

The problem is clearly stated by a 20 year old from a recent qualitative study of female camouflage [1]

> "The amount of girls that aren't diagnosed because they are more likely to camouflage than boys is really bad. I went for so long without being diagnosed because they didn't know that I could pretend to be normal!"

The problem to be solved is the true number of girls with autistic spectrum disorder (ASD). The solution lies in categorizing the data of 1711 children I have managed with ASD with information provided in the detailed histories from the children and their parents.

2 The structure of the study.

The study is inductive [1] using qualitative and quantitative information. The seminal observation is that the carer's second child is easier to recognize than the first. There is no prior hypothesis other than the numerical solution should lie somewhere in the data. The method is analogous to diagnosing a patient but of a population rather than an individual. Like any clinical diagnosis the most important clues are going to be in a detailed history. The process is then initially to seek patterns in the histories and enable the children to be categorized into groups such that the problem can be addressed.

Once patterns became evident it was realized that clues were already present in published data and these were interpreted to triangulate the study results. Like a clinical diagnosis the study uses cumulative qualitative and quantitative information and is necessarily Bayesian in philosophy [2]. Intermediate hypotheses were generated and the supporting information for them was derived from the clinical database and descriptive and hypothesis testing statistics were applied as needed to validate results. The statistics sources are listed in appendix D. Final outcomes of interest were derived both

from the database alone or in combination with data from published studies and equations were derived to validate these results.

3 Study environment.

I have a paediatric private practice focused on behaviour based in South East Queensland. It is bulk-billed under Medicare (service free at point of delivery) and serves primarily the Sunshine Coast and adjacent areas with a drainage population of about 800,000. It covers the full range of socioeconomic status groups and is likely representative of the Australian population as a whole.

The study group was a total of 1711 children 1-18 years of age with ASD diagnosed and/or managed by me between October 2014 and April 2020. Diagnosis was by DSM-5 clinical criteria. The study was stopped when procedural changes in Queensland and Australia, combined with the pandemic, made it likely that I was no longer seeing a representative population.

In Queensland at that time the formal diagnosis was made by a specialist paediatrician or child psychiatrist, with or without advice from allied health professionals (AHPs), but commonly assisted by a psychologist in particular. In specialist private practice all patients had been referred by a general practitioner (family physician) and may have also been seen by AHPs and often provisionally diagnosed prior to referral. I had formal AHP assistance with half my patients. With at least one gatekeeper (GP) and often another (AHP) the important DSM-5 criterion that there was a clinical problem was satisfied.

METHODS, DEFINITIONS AND INTERMEDIATE RESULTS

1 Male/female odds ratio: definitions.

The male/female odds ratio (MFOR) may be defined in various ways. It is the male to female ratio among those with ASD, controlling for the male to female ratio among those in the population of interest without ASD [3]. If we start with a population (ASD and not ASD) of equal numbers of males and females the MFOR is the probability of diagnosing a male with ASD divided by the probability of diagnosing a female. It measures the relative prevalence of boys and girls with ASD in their respective populations. It is a ratio of proportions and mathematically an odds. At birth and during the age range of this study (1-18 years) there is a population excess of males of around 5%. The values for subsets of this age range were found from Australian Bureau of Statistics data from the 2016 census and the number of internal siblings (the key category for the MFOR calculation and defined below) in each subset was counted and the weighted average male/female ratio was found to be 1.055. This correction was applied as appropriate to the unadjusted case numbers. An MFOR using unadjusted case numbers was designated a uMFOR. The most commonly reported MFOR is around 4:1 [4] (p. 57) though it is believed it is probably closer to 3:1 [3]. It is believed there are remaining biases against females though these have not been quantified or even clearly defined. I dealt with gender fluidity by using the sex of assignation at the time of diagnosis. There were only a handful of internal sibling cases and they flowed both ways.

2 Defining the biases and patient categories.

Initially the family knows there is something odd going on with their child's behaviour but has no prior experience to explain it. The behaviour is evident in family settings and/or school. After a variable

number of iterations with various professionals a diagnosis of ASD is eventually made. After the diagnosis is made caregiver knowledge of what to look for usually increases rapidly.

We will consider two biases. The first is that of recognition which is the set of factors preventing the girl reaching the door of the diagnostician's clinic. The second is that of diagnosis which is the set of factors after that point preventing the definitive diagnosis being made. A bias was quantified as the ratio of an MFOR with the bias to an MFOR without it.

The study assumed that the MFOR in simplex families (one child with ASD) would be about the same as for the first child diagnosed in multiplex families (two or more children with ASD). The first child diagnosed in the multiplex family was designated the proband. When simplex and multiplex proband cases were combined or interchangeable they were termed singletons. Subsequent children in multiplex families were designated siblings and would have a lower MFOR due to increased caregiver awareness. Crucially these children are not necessarily younger siblings, because older siblings may have been missed. It is the order of diagnosis that is critical. Half siblings, step siblings or foster siblings in the same household as the proband when I did diagnose them were counted since the key factor is increased awareness after proband diagnosis, not genetic relatedness. There were then three family categories: simplex, proband and sibling. The ratio of a singleton MFOR and sibling MFOR in the family diagnosis category (fig 1) was defined as a recognition bias Br.

All subjects were either diagnosed by me or if already diagnosed when referred for management were assessed by me as meeting DSM-5 criteria. There is variation between clinicians in how these criteria are interpreted so the diagnostic variability in finding the unbiased MFOR was minimised by not including siblings I did not personally diagnose. There were two diagnostic categories. Those designated external diagnosis were referred to me already diagnosed, normally formally by a paediatrician or child psychiatrist and corroborated by me on first consultation. Those referred for confirmation of suspected ASD, a behaviour problem or other diagnosis who on my assessment met DSM-5 criteria for ASD were designated internal diagnosis. A diagnosis bias Bd was defined as the ratio of the MFORs of an external and internal family category (fig 1). If there is residual bias in my diagnosis of girls in either direction by definition I am unaware of it but the internal cases were the only baseline available to me and the external validity of this categorisation will be explored.

During the study period when all cases were diagnosed by a paediatrician or child psychiatrist I was on the watershed where it was possible to categorise all patients to determine the biases from suspicion to final diagnosis. The product of recognition bias and diagnosis bias was defined as ascertainment bias. It will be necessary to assess whether my procedure was externally valid, but at least it meant the children were assessed in a consistent manner by a single clinician. The first aim of this paper is to derive the unbiased MFOR from the dataset and validate it externally. The sibling data are shown in table 1. The categories are shown in fig 1 and we will examine the internal sibling category in detail.
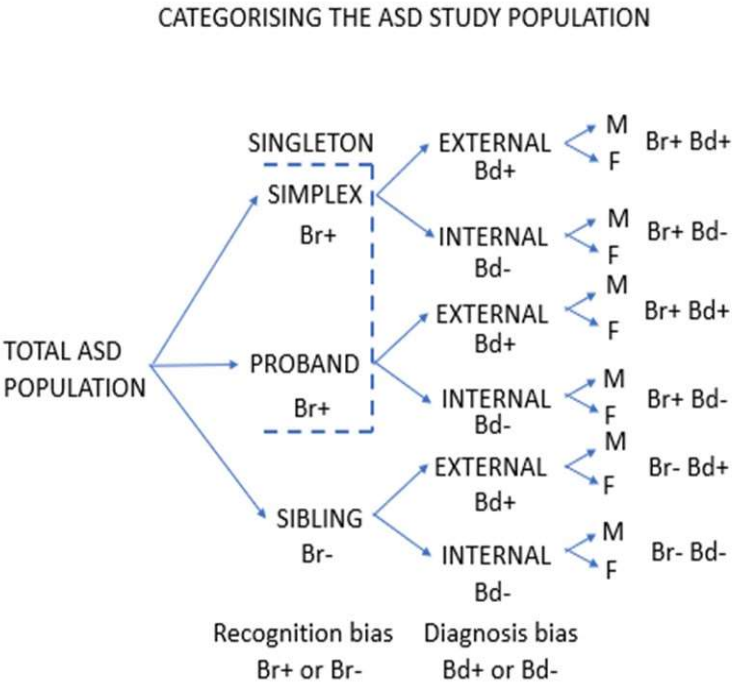
CATEGORISING THE ASD STUDY POPULATION

Fig 1

Table 1: Data for sibling categories

Diagnosis	Male	Female	uMFOR	MFOR
External	44	39	1.128	1.069
Internal	200	253	0.7905	0.7493
Doctor alone	95	127	0.748	0.709
Doctor+AHP	105	126	0.883	0.790
AHP 1st	75	98	0.765	0.725
Doctor 1st	30	28	1.071	1.016

3 Male/Female Odds Ratio: results.

The only category with minimum bias is the internal siblings and the unbiased male/female odds ratio (MFOR) is then 200/(253 x 1.055) ie 0.7493 or 3:4. Statistical confidence boundaries are as follows:

The 95% confidence interval is 0.621-0.900 and the 99% confidence interval is 0.585-0.954.
There is a 95% probability the MFOR is < 0.873 and a 99% probability the MFOR is < 0.932.

This MFOR is much lower than current estimates and we must rule out possible biases in the methodology, explain any data anomalies and seek validation from independent data. First there is the comparison with the external siblings. The external MFOR is 1.069 and the external/internal MFOR ratio 1.426 which measures the diagnosis bias between the groups. The external sibling case numbers are not large but there is a 99% probability the external MFOR is < 1.92 compared to the current suggested

lower limit of 3 [3]. The external and internal ratios are not significantly different: χ^2 for comparison of proportions gives P = 0.1366. This is an issue for calculating the diagnosis bias which will be dealt with later. It does show my diagnostic process is not very different from my medical peers and gives very different results to the current accepted values. Comparison with my AHP colleagues for diagnosis of internal siblings appears to show I am an outlier at MFOR 0.709 but when compared to the MFOR of 0.725 for the group where 21 different AHPs had made the provisional diagnosis there was little difference. χ^2 for comparison of proportions gives P = 0.9113. So I examined the diagnostic pathway in detail.

4 Internal sibling diagnostic pathway.

My practice was to make the diagnosis myself if sure, but to refer the younger children to a psychologist for assessment if I was unsure. The third category had already been assessed by an AHP when referred to me. I looked at the referral patterns for males and females <3, 3 +4 and 5+6 years of age (table 2).

Table 2: Percentages of categorized females: by age {columns} and diagnostic pathway (rows).

Category	<3 years	3+4 years	5+6 years
Doctor alone	{19} (13.7)	{48} (34.5)	{72} (51.2)
Doctor first	{22} (66.7)	{11} (33.3)	0
AHP first	{59} (46.1)	{41} (32.0)	{28} (21.9)

I was indeed seeking diagnostic help for the youngest girls, but this pattern was being repeated by all carers or clinicians, getting AHP testing before medical referral. The MFOR for <7 years was 0.853, reflecting the fact boys get recognized earlier, and MFOR for my diagnosis alone <7 years was 0.874 compared to the AHP assisted MFOR of 0.835 showing I was not biased towards the younger girls compared to assessments by or with AHPs. The median age of diagnosis for girls was 7y1m and for boys 6y7m (appendix 3) and the overall MFOR for siblings 7 years and over was 0.665. The conclusions were all assessors were more comfortable diagnosing girls as they got older, the proportion of girls diagnosed was higher, and I was seeing proportionately more girls directly. My apparent diagnostic biases were functions of age-related referral patterns and not significantly different diagnostic practice.

5 Are the sibling genders found truly random?

The key assumption was that the siblings would be recognized and diagnosed according to the uMFOR. There is one specific situation where this might not be true. If there are two gender discordant children, both undiagnosed, there is going to be a bias towards diagnosing the boy first. The particular order of concern is if the younger child is male and diagnosed first then there is likely to be an excess of male probands and female siblings due to recognition bias in those designated siblings. It was uncertain how to interpret a result if the proband was external because of the effect of diagnosis bias, and I looked at the situation where the first diagnosed was younger than the second in sibships of any size where the proband and first sib were internal and gender discordant. There were 11 younger female probands (38% of total) and 24 younger male probands (35% of total) giving 13 excess female siblings. The difference was not a surprise, but the fact that it is as common for female probands to have older brothers as the reverse was not expected.

When we look at a cross-sectional sample of the population at different ages as we are here for ASD siblings and adjust for excess males, it is a proxy for an assumed cohort where we start with equal gender numbers and assume that with no bias boys and girls with ASD are equally likely to be diagnosed by an arbitrary cut-off, here 18 years. There is however a lag at the older end, because in this group there is a delay in diagnosing girls. The number of girls in the cohort missed after age 18 compared to the boys can be estimated. The 90th centile for diagnosing girls was 14y5m, 1y4m after the boys (appendix C). Over the last 3 years (15-18) 9,7 and 8 girls were diagnosed averaging 8/year so we can say about 8 x 4/3 = 10.7 extra girls should be counted. The difference between excess girls found in the sibling dyad and cohort girls missed is 2 girls. There are also the girls who are never found since there is likely to be incomplete recognition of uncertain magnitude before 18 years. Another way of assessing whether the unknowns cancel, in particular male bias in the initial younger brother/older sister dyad, is to compare the MFOR of siblings in families where there is a proband/sibling pair only with the MFOR of the balance of the internal siblings from larger families where the recognition should be greater. These were 117/(148 x 1.055) and 83/(105 x 1.055), giving 0.7493 and 0.7493 respectively. This was a numerical fluke but does suggest the overall MFOR result of 0.7493 is a reliable estimate.

6 Translating the MFOR to numbers of practical importance.

Finding the MFOR does not directly lead to the true female prevalence or estimate the proportion of girls missed. These can be derived from published data combined with the MFOR or using Bayes' theorem where the true ASD prevalence can be found using a more visible comorbidity as a tag. If we know the proportion of the tag in ASD and conversely the proportion of ASD in the tag plus the prevalence of the tag condition we can use Bayes' theorem to find the prevalence of ASD. Bayes' theorem is integral to clinical diagnosis though we often do not realise we are using it [5]. There is argument about the precise definition of many psychiatric conditions but this does not invalidate the calculation if the conditional probabilities employ similar definitions.

The conditional probabilities and prevalence are available for borderline personality disorder (BPD). Data are available for lifetime prevalence of BPD in women, 6.2% [6] and adult data which should reflect lifetime prevalence for BPD in ASD, 6 of 40 = 15% [7] and ASD in BPD, 6 of 41 = 14.6% [8]. ASD is of course a lifelong condition. The conditional datasets are unfortunately small but they are all we have and we shall see where they lead. By Bayes' theorem:

$$P (ASD|BPD) . P (BPD) = P (BPD|ASD) . P (ASD)$$

$$0.146 \times 0.062 = 0.15 \times P (ASD)$$

$$P (ASD) = 0.060$$

Venn diagram of Bayes Theorem

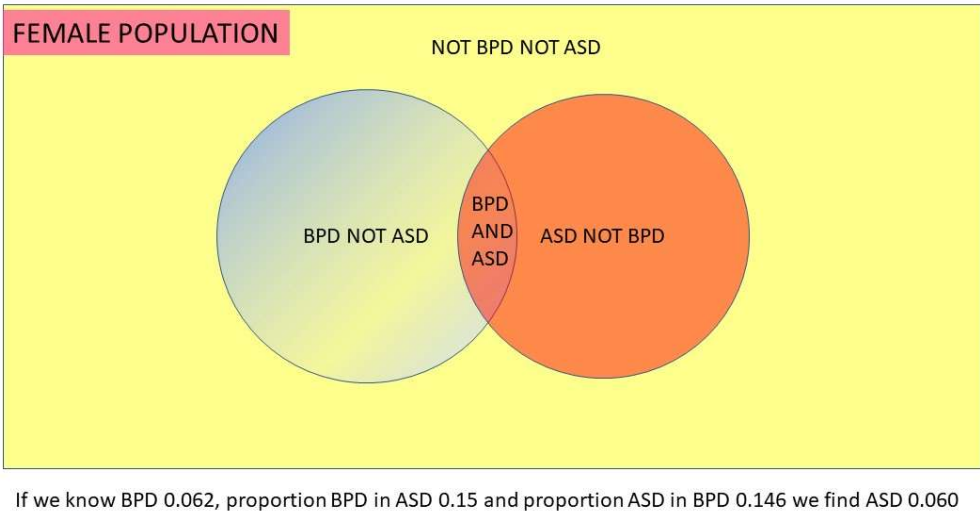If we know BPD 0.062, proportion BPD in ASD 0.15 and proportion ASD in BPD 0.146 we find ASD 0.060

Fig 2

This indicates 6.0% = 1/16.7 women have ASD (fig 2). A recent US prevalence [9] estimate of 1.25% is 1/80 and 1/16.7 is 4.8/80. If we find 1 and miss 3.8 we are missing 3.8/4.8 x 100% or 79% of the girls. The same paper showed an overall ASD prevalence in the US from 2014 to 2016 of 2.47%. There was no statistical increase over that time though the number was increasing slightly. This may be a true number change. As knowledge of ASD has increased the prevalence will have followed a sigmoid learning curve, and while it is certainly flattening the asymptote has probably not yet been reached (shown in stylised form in fig 3.)

This asymptotic approach has different implications for each gender. For males it suggests the limit of detection by DSM-5 is being approached. For females it suggests no major advance in reducing bias has occurred during the study period.
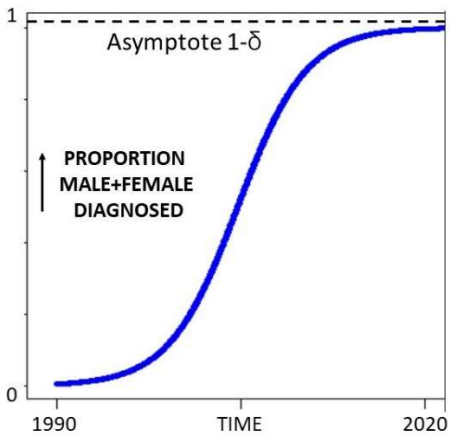

Fig 3.

The overall prevalence in 2016 was 2.76%. The male prevalence over the 3 years was 3.63% and so the likely male prevalence in 2016 would be (2.76 x 3.63)/2.47 = 4.06% and it may have a little way to go to truly flatten with diagnostic saturation. Applying the MFOR I found of 0.7493:1 the corresponding female prevalence would be 4.06/0.7493 % ie 5.42% or 1 in 18.5 girls. This translates to missing 77% of the girls.

The ASD prevalence estimates are strikingly divergent from current estimates but an independent tag and published data have given similar results congruent with my finding. The result using my data is derived from the MFOR and is relative to the diagnostic rate for males but the Bayesian result is entirely independent and depends only on data for females.

We have now derived the male/female odds ratio (MFOR) and combined with published data have found that nearly 80% of girls with ASD are missed. We will now use the entire patient database alone to estimate the biases and the proportion of girls missed and derive an equation to model the current published MFOR values. We will also use independent information from the patient histories to corroborate the results.

7 The algebra of the biases.

The variables in the final formulae are expressed as probabilities (proportions), odds ratios and biases.

Table 3: Notation scheme for variables

Entity	Subscript	Subscript
B bias	a ascertainment	m male
D population	b sibling	o no bias
N number	d diagnosis	p proband
P proportion	f female	r recognition
R odds ratio	i reciprocal	u unadjusted

The reciprocal of the bias gives the proportion of girls found when the bias operates. It is compared to the males where it is assumed there is no bias and the two groups are starting with equal numbers, having adjusted for the male excess in the general population. Whether we are finding 100% of the males is moot (fig 3) but it must form the baseline for gender comparison. Then if 100% of males are found and the bias is 5, 20% of the females are found and 80% are missed. Crucially recognition bias must always precede diagnosis bias in the ascertainment process.

If recognition bias B_r is present the proportion of girls found is $1/B_r$ and the proportion missed is $1-1/B_r$. It is this proportion $1/B_r$ who will be assessed and subject to diagnosis bias B_d. The proportion finally ascertained with ASD is $1/B_rB_d$. B_rB_d is defined as the ascertainment bias B_a. The population proportion missed on diagnosis is $1/B_r-1/B_rB_d$ which reduces to $(B_d-1)/B_rB_d$ or $(B_d-1)/B_a$. The effect of B_d is always going to be less than B_r for a given value since B_d operates on proportion $1/B_r$ of the population and $1/B_r$ is always <1. If B_r is large the effect of B_d will be small on the population as a whole since the function is hyperbolic. This is shown in fig 4 where 80% of girls are missed if $B_a = 5$ and let $B_r = B_d = 5^{0.5}$.
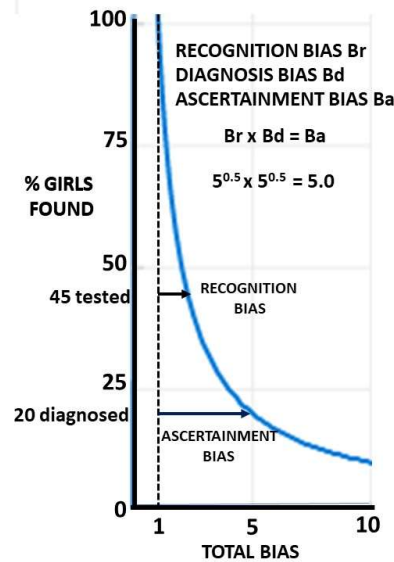
Fig 4

Only 45% of girls are recognized, then 45% of the 45% ie 20% are finally diagnosed. A large recognition bias will make losses on diagnosis small at a population level. The proportions found rapidly drop at first but as the slope flattens they become relatively insensitive to minor changes in the bias. The importance of diagnosis bias depends on your perspective. If you are a girl about to be tested it is very important. If you are a planner deciding where to put funding then improving recognition may well be more important.

8 Diagnosis bias.

There were two in the dataset. The first was the singleton (proband + simplex) bias, the ratio of the external singleton unadjusted MFOR (uMFOR) to the internal singleton uMFOR. The other was the sibling bias, the ratio of the external sibling uMFOR to the internal sibling uMFOR. For biases the factor 1.055 cancels and the actual numbers of cases can be used. Neither bias was statistically significant by χ^2 for comparison of proportions so a meta-analysis was done using Stouffer's method [10] where the sum of the Z scores between the external/internal ratios giving the singleton and sibling biases was divided by $2^{0.5}$. Without this procedure it could be assumed that there was no significant diagnosis bias at all and this would be a serious type 11 error. Z for the singleton ratios was 1.3327 and for the sib ratios was 1.4732. The combined Z was 1.985 giving a P value of 0.0236. During data collection the external MFORs were always greater than the internal ones and we are interested principally in one tail but the 2 tailed P would be 0.0472. The two biases were then combined by a weighted mean into the diagnosis bias Bd:

$[(\sum\text{singletons x ext sing ratio/int sing ratio}) + (\sum\text{siblings x ext sib ratio/int sib ratio})]/\sum \text{singletons + siblings}$

An important inference from the fact both diagnosis biases are small and similar is that the external sibling MFOR after removing the weighted mean diagnosis bias is little different to the key variable the internal sibling MFOR. It is 0.845 compared to 0.749: χ^2 for comparison of proportions gives P = 0.612. This implies siblings are presenting to my medical peers and myself in much the same gender proportion.

9 Correction to an MFOR with recognition bias Br only.

In order to make the true MFOR (designated Ro) as accurate as possible only children I evaluated were counted as probands or simplex. If there were one or more prior diagnosed siblings with ASD the proband or simplex gender ratio is going to be falsely low because a proportion Pb will have prior diagnosed siblings I did not evaluate and will themselves be siblings without recognition bias. The ratio of external siblings to external singletons was used as an estimate for Pb. The ratio of external siblings to external singletons should approximate the proportion of internal singletons who are actually siblings themselves of external probands. An external singleton is about as likely to have an older sib as a younger one so the latter can serve as an estimate of the former. I neglected possible differences between simplex and proband because any numerical effect on the final estimates was likely to be small. The value for Pb was 0.227 and this value derived from externally diagnosed patients was deemed appropriate when estimating recognition bias and the biased MFOR because the external data were generalizable in calculating both these results which reflected external assessments.

I first used the internal proband ratio for calculation for Br because it related directly to the siblings. A case can be made for including the simplex cases as well since there is no obvious biological difference to the probands and so I also used the weighted mean of both groups (internal singleton MFOR) for a separate estimate of Br. This is discussed in detail later.

To derive a value for Br from the probands the internally diagnosed proband MFOR (Rpr) is then corrected for the unrecognized prior siblings and the ratio: (corrected MFOR of internally diagnosed probands with Br)/(MFOR with no Br) will be the true recognition bias Br ie Br = corrected Rpr/Ro

The derivation of the corrected Rpr is shown in appendix A. The formula is:

$$(Rpr-Ro.F)/(1-F) \text{ where } F = Pb.Pf.(Rpr+1)$$

10 External validity of the variables.

Do the variables derived from my data reflect published results? To derive an MFOR including recognition and diagnosis biases we start with a population of ASD children distributed in the proportions of the unbiased MFOR as Ro/1 males to females. This is the algebraic equivalent of the proportion of males Pm to the proportion of females Pf. The population has siblings (broadly defined as with minimum recognition bias) in proportion Pb as derived above. The siblings will be distributed by gender in proportions Pm and Pf. Female siblings of male singletons will be recognised according to diagnosis bias rate Bd. Let its reciprocal be Bdi. Only proportion Bai of the female singletons will be ascertained as having ASD, where Bai is the reciprocal of the ascertainment bias Ba. Their male siblings will be ascertained but their female siblings will be subject to diagnosis bias Bd. The major proportion of the female singletons (1-Bai) will not be ascertained due to ascertainment bias Ba. Their male siblings will be ascertained but their female siblings will be subject to Ba because the female singletons themselves were not ascertained. All the variables in the equation are derived from the unbiased MFOR Ro, the recognition bias Br, the diagnosis bias Bd and the ratio of external siblings to external singletons Pb. These in turn are all derived from the study dataset. The biased MFOR is then the sum of all the boys divided by the sum of all the girls. In appendix B the groups outlined above are shown in table A1 and the biased MFOR is derived. The formula is:

$$Ba.Ro(Pb+1)/(1+Pb.Pf(Ro.Bdi.Ba+Bdi-Bai+1))$$

Now we can calculate the outcome of the biases in the ASD population.

Table 4: Complete dataset of 1711 children for calculations

Category	Diagnosis	Male	Female	uMFOR	MFOR
Simplex	External	142	57	2.491	2.361
	Internal	376	193	1.948	1.847
Multiplex proband	External	119	48	2.479	2.350
	Internal	171	69	2.478	2.349
Sibling	External	44	39	1.128	1.069
	Internal	200	253	0.791	0.7493

The calculation for the multiplex proband + sibling categories (detailed in appendix B) gives:

Biased MFOR = 3.31. % Singleton females missed on: Recognition 79.1 Diagnosis 4.3 Total 83.4

For every 100 singleton boys found we should find 133 girls. We find 22 and miss 111.

For the weighted mean of simplex and proband categories the corresponding results are:

Biased MFOR = 2.76. % Singleton females missed on: Recognition 73.1 Diagnosis 5.7 Total 78.8

For every 100 singleton boys found we should find 133 girls. We find 28 and miss 105.

A recent meta-analysis [3] found that good quality clinical studies yielded a mean MFOR of 3.32 and they concluded the likely value was closer to 3. My results are consistent with these findings.

11 Which Internal Singleton MFOR?

It is not obvious which internal singleton MFOR is the better one to use to calculate the recognition bias. The proband value relates to the actual siblings, but assuming the simplex children are not fundamentally different the weighted mean of proband and simplex is derived from a larger number of cases. The recognition bias is much larger than the diagnosis bias and must precede it in the calculation and the results are where the slope of the hyperbolic function is flattening and so are not very sensitive to variation. For clinical and policy simplicity I rounded Br to 4 and Ba to 5 giving Bd a value of 1.25 (see Summation section for details). The MFOR can be rounded to 0.75 so we then have 4 girls for every 3 boys. We have 75% of girls not recognized, one in five of the 25% assessed are not diagnosed, leading to 80% overall missed (fig 5). If we use these variables in the biased MFOR equation together with a Pb of 0.227, the only variable not derived from Br, Bd and the unbiased MFOR, we find a biased MFOR of 2.88, and for every 100 singleton boys found we miss 107 girls. For comparison the MFOR of the weighted prevalences of 3.63% for males and 1.25% for females in the recent US study[9] is 2.90. The first 3 years of data for my study were collected at the same time as the US data suggesting the study values are accurate and generalizable.

The finding that 20% of assessed girls are missed is highly clinically significant and shows the importance of the diagnosis bias meta-analysis but this is only 5% of the total girls with ASD in the population. Whichever set of these biases is chosen there are overall more singleton girls who are missed than boys who are found.
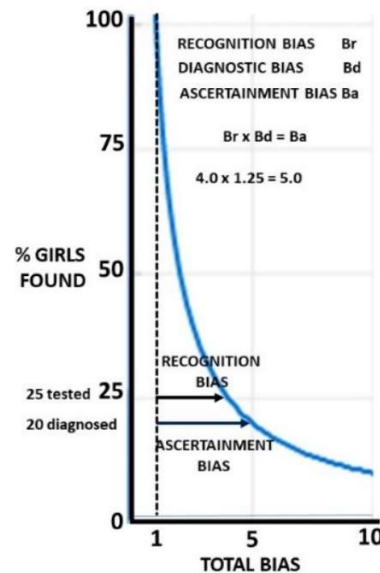
Fig 5

12 Explaining a data anomaly.

A puzzling finding in the database was that the internal simplex MFOR was persistently about 0.5 less than the internal proband MFOR. This was not statistically significant by χ^2 for comparison of proportions but it was stable and the issue was whether the non-significance was a type 11 error and the data was trying to say something. It would require something different happening in each group and one possibility was that different proportions of the singletons were in fact siblings. The dataset records the detailed family relationships. The singletons will be distributed by gender by MFOR (Ro) ie Ro/1. We assume all the males are found. Br is the recognition bias and Pb is the external sibling/singleton ratio describing the proportion of singletons who are actually sibs and do not have recognition bias.

Simplex cases may have at least one prior diagnosed sibling I do not manage. Probands have at least one sibling diagnosed after them that I manage so to have at least one prior diagnosed sibling they must have at least two siblings and so the proportion must be less than Pb because fewer families have larger numbers of children. I tallied the total number of families with a proband and one internal sibling (206) and the number with a proband and at least two internal siblings (120). Then the correction factor for Pb for the proband category will be the proportion of families with one or more sibs who have more than one sib which is 120/326 ie 0.368.

The proportion of females recognized is 1 x 1/Br. The proportion who are siblings is Pb x 1/Br. But this is not relevant in this group because they are included in the proportion 1/Br who have all been found. In the group not recognised (1-1/Br) there are Pb (1-1/Br) siblings who as siblings will be found. The total proportion of females found is then 1/Br + Pb(1-1/Br) and the singleton MFOR is Ro/[1/Br + Pb(1-1/Br)] which simplifies to:

$$RoBr/[Pb(Br-1)+1]$$

The proband MFOR formula is then:

$$RoBr/[0.368Pb(Br-1)+1]$$

If we use the simplified Br of 4 and MFOR of 0.75 in the singleton equation we find MFOR 1.785 with the actual value 1.979. In the proband equation we find 2.399 with the actual value 2.349. The first derived value is 10.9% high, the second 2.1% high. The reasonably close match of model and data implies sibling recognition accounts for most of the anomaly.

13 Corroboration of the recognition bias.

Asking about camouflaging is part of my standard clinical history and I tallied the experience of the families of 100 school age diagnosed girls in consecutive clinics in my practice. The criterion for camouflage used was being Ms Jekyll at school and Ms Hyde at home, with the transition at the school gate:

<p align="center">"I fall to pieces." [1]</p>

<p align="center">"I was unbearable with my mother, but at school I was perfect." [11]</p>

The study was effectively random, had a 100% response and 88 girls behaved in this way. This is consistent with the histories of adult women where 93% had camouflaged[1]. Part way into the study period I added the clinical question as to whether school observations had helped with the diagnosis and had a 100% response from 69 families with school recognition failure of 72%. I later realised this was a far more important question for diagnosing girls since the 88% is the rate of the behaviour and the 72% is the rate of failure to recognise the behaviour at school. Social camouflaging at school is not just hiding ASD from the class teacher, but also from the special needs teacher and the guidance officer and 72% of families described school observations as having not been of use in initial diagnosis. I did not extend the number of responses because by not recognising the importance of this variable when getting the histories I was not unconsciously biasing the responses and affecting the accuracy. I did continue to ask the question as part of the normal diagnostic history for both girls and boys.

 The two settings where children are observed closely over time are school and home, and individual families without a known child with ASD are unlikely to be as skilled in recognition as school officers charged with dealing with these sorts of problems. If we then assume those who are not recognised at all camouflage in general better than those who are recognised by an agency other than school then the proportion of those diagnosed who are missed at school of 72% sets a lower bound on the proportion of total females with ASD not recognised. Then the proportion not recognised will be 72+δ % which is an independent corroboration of the recognition bias using information about female behaviour only.

14 Aftermath.

In early 2020 there were 2 major changes to ASD diagnosis in Queensland. State funding of school assistance no longer required medical sign off and the National Disability Insurance Agency did not require medical sign off for non-medical assistance for ASD. In addition medical services were no longer face to face because of the pandemic and it was unclear what the overall covid effect would be. We will examine the effect on the data to see if terminating the study at that point was justified.

The only remaining absolute requirement for medical input was the prescription of medication for the comorbid conditions of ADHD, sleep problems, anxiety and angry, irritable or violent behaviour. This would inevitably skew the referrals to boys and to the singleton girls who were in the category already recognized due basically to behaving like the boys. Siblings would span the range of behavioural severity but it was thought likely that in families with prior experience of managing ASD the better behaved girls would possibly have their assessment held back due to covid and then be less likely to need medical

referral anyway. The result would be that I would see a skewed population with fewer girls, in particular sibling girls, and the MFOR would be factitiously raised. The final dataset of 2246 children up to 020721 is presented below.

Table 5: Final dataset of 2246 children.

Category	Diagnosis	Male	Female	uMFOR	MFOR
Simplex	External	192	76	2.526	2.395
	Internal	512	251	2.040	1.933
Multiplex proband	External	136	59	2.305	2.185
	Internal	224	101	2.218	2.102

Sibling diagnosis	Male	Female	uMFOR	MFOR
External	49	48	1.021	0.968
Internal	282	316	0.8924	0.8459
Doctor alone	150	169	0.888	0.841
Doctor+AHP	132	147	0.898	0.851

We will first examine the entire internal category and compare internal singletons and sibs.

Table 6: Internal group.

Time of data collection	Total uMFOR	Sibling singleton ratio
Up to 170420	1.450	0.560
Up to 020721	1.524	0.550
180420 to 020721	1.771	0.520

These results show that overall both fewer girls and fewer siblings were being diagnosed. Within the sibling category the MFOR had risen. The MFOR up to 170421 was 0.7493. The MFOR from 180420 to 020721 was 1.302. A χ^2 for comparison of proportions gave P = 0.0090, so clearly as predicted female sibs were not being referred for diagnosis. The proportion of siblings referred for diagnosis having seen an AHP first prior to 170420 was 38.2%. The proportion after this date was 26.2%. The P value by χ^2 was 0.0085, supporting the hypothesis that they would not be referred unless a doctor was needed. The ratios for sibs who saw either an AHP or myself first both rose, suggesting that fewer girl sibs were presenting to anyone for diagnosis. This was most likely due to covid. Overall there was a major difference in referral patterns from early 2020 justifying termination of the study. It also demonstrates the serious distortion of results when the sample population is not truly representative.

15 Summation.

This paper has used a total of 4 methods to determine the recognition and/or ascertainment biases. The method using only the clinical database has arrived at 2 possible sets of values for each. A value of the ascertainment bias has been derived from the combination of the data derived MFOR and a recent estimate of female prevalence in the US. Another value has been derived by entirely independent

published data on the relation between ASD and borderline personality disorder using Bayes' Theorem. An estimate of recognition bias has been made from recognition failure at school using the clinical histories of my patients. The mean of each set of estimates was then found. This ensemble averaging is used for estimating the tracks of tropical cyclones [12]. The values were not weighted since there was no evidence one estimate was better than another. The final working values were rounded to the nearest whole percentage (Table 7). The characteristics of the biases mean that estimates of recognition bias of 4 and ascertainment bias of 5 are sufficiently accurate for clinical and planning purposes.

Table 7: Pathways to finding girls missed.

Information source	Not recognized %	Not ascertained %
Study + US prevalence		77
Borderline personality		79
School recognition	72+δ	
Study alone	73.1, 79.1	78.8, 83.4
Mean, working value and bias	74.7+δ/3, 75, 4	79.6, 80, 5

DISCUSSION

1 Outcome of the biases and future directions.

Currently while diagnosis bias is important for individual patients the bigger problem for all girls with ASD is the lack of initial recognition. If diagnosis were improve, so half those currently missed were found there would on a population basis be only a 2.5% improvement. Due to the hyperbolic nature of the bias relationship a significant advance in recognition would rapidly lead to diagnosis bias becoming quantitatively more important as the values move to the left on the hyperbola. If the recognition bias were halved those missed on diagnosis would double to 10% of the ASD population without any change in diagnostic practice. The two problems need to be tackled in tandem to remedy what is currently a very serious failure of clinical intervention.

There is emerging qualitative research from ASD women on their camouflaging practices [1]. This needs to be combined with similar information from the caregivers: front line heath workers who first encounter the problem, teachers, child care workers and above all families including those who have ASD children and those who do not. There will be other reasons for missing girls, but I predict camouflaging will be an upstream causal factor in them all. This information will provide the context for designing programs to better educate all those potentially involved in finding girls with ASD.

2 Hypothesis.

Is there evidence for a reason why girls might outnumber boys? The aetiology of ASD is largely but not entirely genetic [13] and there may well be a cultural component. Section D of the DSM-5 criteria [4 p57] requires the entity to cause a clinically significant impairment. There is abundant qualitative evidence of the distress girls with ASD suffer growing up [1,11,14]. As a clinician it is clear that the central problem of ASD is poor reciprocal communication and the social communication expectations for girls are higher than for boys. Studies showing gendered genetic differences have a problem. For any study to demonstrate a true quantitative gender difference in any characteristic such as gene distribution the male and female populations studied which are the denominators of the proportions of

the characteristic compared must not have different degrees of ascertainment bias. Unless the gender samples are truly representative the proportions of the characteristic cannot be compared [14]. There are at least 102 genes to distribute [15] suggesting the most economical theoretical genetic contribution to the MFOR is 1:1. The cultural pressure leading to a diagnosable disorder will then tip the balance to females. Camouflaging appears to begin very early and there must be environmental factors with networked causal pathways for camouflaging and other factors but here I frame no hypotheses.

3 Why does this matter?

From these results it appears for every 1000 women about 60 have clinical ASD. By 18 years of age 12 have been diagnosed and 48 have not. If this is really true of ASD in women then it is a significant upstream factor in female mental health. Childhood and adolescence are very difficult for girls with undiagnosed ASD and constant psychic trauma is inevitable including vulnerability to sexual exploitation [11]. I do not diagnose adults but from my experience with the histories of diagnosed or probably affected mothers, which are common due to the high heritability [13], anxiety is very common in ASD and adult women describe relentless mental trauma from a young age with no cause found, or worse, a whole gamut of incorrect or incomplete causes. It appears that women with ASD get an alphabet soup of diagnoses including borderline personality disorder, eating disorders, bipolar disorder, schizoaffective disorder, schizophrenia, post-traumatic stress disorder, sensory processing disorder, intermittent explosive disorder and adult ADHD as well as the varieties of anxiety, agoraphobia, panic disorder and depression, serially and together. I advise on a pathway to diagnosis for a disturbing number of mothers. ASD is not a mental illness and an individual may have one or more of these conditions as comorbidities, but without the upstream causal factor of ASD identified she will never fully get to grips with her condition and gain psychic relief by understanding herself.

4 Autistic Spectrum Disorder and Autistic Spectrum Condition (ASC).

There is a view that the term ASD is stigmatizing and ASC will serve for both the strengths and difficulties of those on the spectrum [1]. I do not think ambiguity is helpful in a clinical discussion. I believe each term has value in context. As a clinician I only make a diagnosis if there is a disorder. Many of those with features of the spectrum adapt without clinical intervention. An ASC is not a disease to "cure" or a disability to treat. For those whose minds are disordered the diagnosis is *the* fork in the road. They then have the opportunity through self-understanding and therapeutic intervention to transition to an ASC. This is the satisfactory end point. The other tine of the fork leads to continuing disorder or descends into mental illness. From my observations the perceived stigma of a diagnosis is definitely a feature of both recognition and diagnosis bias. I believe clear differentiation between ASD and ASC and seeing diagnosis as the key to enabling the transition will reduce bias, in particular for girls.

CONCLUSION.

From a practical perspective the results of this study need only be sufficiently precise for informed decision making on diagnosis, management and service development. This is provided by:

 # three variables, the male/female odds ratio, the recognition bias and the diagnosis bias which are described by three numbers 3/4, 4, 5/4.

three rules: As biases occur their values multiply, the proportion of girls found is the reciprocal of the bias and *above all* attend to mother's history.

APPENDICES.

Appendix A. Derivation of ratio of internally diagnosed probands who are true singletons, corrected Rpr.

For a total number of singletons in ratio N_m/N_f we have number of males N_m adjusted for the demographic excess of males by 1.055 and females N_f. The proportion of cases with low recognition bias is P_b. These cases are effectively sibs and males and females must be removed from the male and female singleton numbers in the proportion of the MFOR R_o. These proportions are $R_o/(1+R_o)$ for males and $1/(1+R_o)$ for females. With overall proportion P_b to be removed and total number of cases N_m+N_f the actual male and female numbers of siblings will be $P_b.(N_m+N_f).R_o/(1+R_o)$ and $P_b.(N_m+N_f)/(1+R_o)$ respectively. This is sibling proportion x total cases x proportion male or female.

Then the numbers of true singletons will be:

Male: $N_m - P_b.(N_m+N_f).R_o/(1+R_o)$ and females: $N_f - P_b.(N_m+N_f)/(1+R_o)$ respectively.

Divide throughout by N_f and substitute P_f (proportion of females in R_o) for $1/(1+R_o)$:

Male: $Rpr - P_b.(Rpr+1).R_o.P_f$ and females: $1 - P_b.(Rpr+1).P_f$ respectively.

Let common factor $P_b.P_f.(Rpr+1)$ be F. Then corrected biased gender ratio with siblings removed:

$$Corrected\ Rpr = (Rpr-R_o.F)/(1-F)$$

Appendix B. Derivation of biased MFOR.

The categories are summarised in table A1. All categories except {not ascertained} female singletons are counted. Each sibling category is: ASD population gender proportion $[R_o/1 = P_m/P_f]$ x sibling proportion x sibling gender proportion x proportion found with bias operating.

Table A1.

Categories	Singleton proportion	Male sib proportion	Female sib proportion
Male singletons Ascertained siblings	Ro	Ro.Pb.Pm.1	Ro.Pb.Pf.Bdi
Ascertained female singletons Ascertained siblings	1.Bai	Bai.Pb.Pm.1	Bai.Pb.Pf.Bdi
{Not ascertained} female singletons Ascertained siblings	{1-Bai}	(1-Bai)Pb.Pm.1	(1-Bai)Pb.Pf.Bai

The biased MFOR is the ratio of the proportions of all the boys and all the girls who have been found:

$$(Ro+Ro.Pb.Pm+Bai.Pb.Pm+(1-Bai)Pb.Pm)/(Bai+ Ro.Pb.Pf.Bdi+Bai.Pb.Pf.Bdi+(1-Bai)Pb.Pf.Bai)$$

Simplify:

$$Ro+Pb.Pm.(Ro+1)/(Bai+Pb.Pf(Bai.Bdi+Ro.Bdi+Bai(1-Bai))$$

Now Pm = Ro/(Ro+1)

Then numerator is Ro+ Pb.Ro(Ro+1)/(Ro+1)

Ro+Pb.Ro

Ro(Pb+1)

Multiply numerator and denominator by Ba giving biased MFOR:

$$Ba.Ro(Pb+1)/(1+Pb.Pf(Ro.Bdi.Ba+Bdi-Bai+1))$$

Calculation of biased MFOR for multiplex probands + siblings (Ro = 0.7493):

Weighted diagnosis bias Bd:

[(∑sing x Ext sing ratio/int sing ratio) + (∑sib x Ext sib ratio/int sib ratio)]/ ∑ patients

Bd = 1.265   Reciprocal weighted diagnosis bias Bdi = 0.791

Pf female proportion 1/(1+Ro): 0.5717   Pb Ext Sibling/singleton ratio: 0.227

Rpr: 2.349 Corrected Rpr = (Rpr-Ro.F)/(1-F) where F = Pb.Pf(Rpr+1)

(2.349 – 0.7493 x 0.4346)/ 0.5654   F = 0.227 x 0.5717 x 3.349 = 0.4346

Corrected Rpr 3.579/Ro: 0.7493 = Br: 4.7759 x Bd: 1.265 = Ba: 6.0415 Bai = reciprocal Ba: 0.166

Biased MFOR = Ba.Ro(Pb+1)/1+Pb.Pf(Ro.Bdi.Ba+Bdi-Bai+1)

5.5545/1.6756

Biased MFOR = 3.31

Appendix C.   Table A2: Centiles for sibling age at diagnosis.

Centile	10	25	50	75	90
Boys	2y8m	3y10m	6y7m	9y6m	13y1m
Girls	3y0m	4y1m	7y1m	10y7m	14y5m

Appendix D. Statistics websites.

Population proportion confidence interval.

https://select-statistics.co.uk/calculators/confidence-interval

Comparison of proportions.

https://www.medcalc.org/calc/comparison_of_proportions.php

Z Score Calculator for 2 Population Proportions.

www.socscistatistics.com › tests › ztest

REFERENCES.

1 Hull, L., Petrides, K.V., Allison, C. et al. "Putting on My Best Normal": Social Camouflaging in Adults with Autism Spectrum Conditions. *J. Autism Dev. Disord.* **2017**; *47*:2519–2534.

2 Gill, C.J., Sabin L., Schmid, C.H. Why clinicians are natural Bayesians. *B.M.J.* **2005**; *330*:1080-1083.

3 Loomes, R., Hull, L., Mandy, W.P.L. What is the Male-to-Female Ratio in Autism Spectrum Disorder? A Systematic Review and Meta-Analysis. *J. Am. Acad. Child Adolesc. Psychiatry* **2017**; *56*:466-474.

4 American Psychiatric Association. *Diagnostic and Statistical Manual*. 5th ed. Arlington, VA: American Psychiatric Association;2013; p54.

5 McCrossin, R. Clinicians have to be Bayesians. *B.M.J.* **2005**; *330*:1390–1391.

6 Grant, B.F., Chou, S.P., Goldstein, R.B. et al. Prevalence, Correlates, Disability, and Comorbidity of DSM-IV Borderline Personality Disorder: Results from the Wave 2 National Epidemiologic Survey on Alcohol and Related Conditions. *Clin. Psychiatry.* **2008**; *69*:533–545.

7 Hofvander, B., Delorme, R., Chaste, P. et al. Psychiatric and psychosocial problems in adults with normal-intelligence autism spectrum disorders. *B.M.C. Psychiatry.* **2009**; *9*:35.

8 Ryden, G., Ryden, E., Jetta, J. Borderline Personality Disorder and Autistic Spectrum Disorder in Females-A Cross Sectional Study. *Clin. Neuropsychiatry.* **2008**; *5*:22-30.

9 Xu, G., Strathearn, L., Liu, B. et al. Prevalence of Autism Spectrum Disorder Among US Children and Adolescents, 2014-2016. *J.A.M.A.* **2018**; *319*:81-82.

10 Whitlock, M.C. Combining Probability from Independent Tests: the weighted Z-method is superior to Fisher's approach. *J Evol. Biol.* **2005**; *18*:1366-1373.

11 Bargiela, S., Steward. R., Mandy, W. The Experiences of Late-diagnosed Women with Autism Spectrum Conditions: An Investigation of the Female Autism Phenotype. *J. Autism Dev. Disord.* **2016**; *46*:3281–3294.

12 Jun S, Kang N-Y, Lee W et al. Multimodal Ensemble Forecast for Tropical Cyclone Tracks in the Western North Pacific. *Atmosphere.* **2017**:*8*:174 https://doi.org/10.3390/atmos8090174 .

13 Bai, D., Yip, B..HK., Windham, G.C. et al. Association of Genetic and Environmental Factors with Autism in a 5-Country Cohort. J.A.M.A. Psychiatry **2019**; *76*:1035-1043.

14 Hull, L., Petrides, K.V., Mandy, W. The Female Autism Phenotype and Camouflaging: a Narrative Review. *Rev. J. Autism Dev. Disord.* **2020**; *7*:306-317.

15 Sallerstrom, F.K., Kosmicki, J.A., Wang, J. et al. Large-scale Exome Sequencing Study Implicates Both Developmental and Functional Changes in the Neurobiology of Autism. *Cell.* **2020**; *180*:568-584.

BIOGRAPHY

Dr McCrossin graduated from Sydney University MB BS (Hons 1) in 1971. He has a Masters of Public Health from the University of Queensland and is a Fellow of the Royal Australian College of Physicians. He has had a dual career as a paediatric clinician and hospital administrator. He was chief resident at the Royal Alexandra Hospital for Children in Sydney and medical superintendent of the Royal Children's Hospital in Brisbane. His clinical career has been as a specialist in general paediatrics and endocrinology, particularly in paediatric diabetes. Both careers made him acutely aware of the importance of behaviour and, reasoning that managing paediatric behaviour problems could surely not be more difficult than managing teens with diabetes or hospital consultants, he has in semiretirement focused on autism and in particular on the diagnosis and management of girls on the spectrum.

COVER LETTER

It is generally accepted in the Autism community, including clinicians, parents and patients, that girls are being significantly underdiagnosed, with their gender specific behaviour leading to biases causing under recognition in the population and underdiagnosis by professionals. Qualitative research shows this becoming particularly problematic for adolescents and associated with the development of serious mental illnesses. The biases have not been clearly defined and there has been no attempt at quantifying them at all. This paper has attempted to define and quantify the biases. The result was that the problem is much more common that expected. If true it is a significant upstream cause of female mental illness where if not picked up becomes severely exacerbated in the adolescent years. If true it affects tens of millions across the globe, and if diagnosed there are very effective interventions. It concerns clinicians, children, adolescents and families and thus fits the scope of the journal and the special issue in particular.