Volume One

Research Component

A developmental scale of early social cognition in autism spectrum disorder

by

Philippa von Rabenau Lewington

A thesis submitted to the University of Birmingham for the degree of Doctor of Clinical Psychology

Department of Clinical Psychology

School of Psychology

The University of Birmingham

May 2018

Thesis Overview

Volume One

Volume One presents three papers. The first paper reports a meta-analysis exploring the prevalence of autism spectrum disorder (ASD) in rare genetic syndromes. Further analyses consider sources of variance in the data, evaluating the influence of methodological factors and sample characteristics on ASD prevalence estimates. The second paper is an empirical study examining the developmental trajectory of early social cognitive skills in children with ASD. The third provides a summary of the reported research, for dissemination to research participants and professionals in clinical and educational services.

Volume Two

The second volume consists of five Clinical Practice Reports (CPRs). CPR one presents two formulations, using cognitive-behavioural and systemic models, exploring anxiety experienced by a 50-year-old man presenting at a community learning disability service. CPR two describes a service evaluation project evaluating staff attitudes towards the use of the 'Friends and Family Test' feedback tool. CPR three presents a case study, reporting a systemic intervention with a 44-year-old woman experiencing low mood. CPR four is a single-case experiment, evaluating the effectiveness of a cognitive-behavioural intervention for self-harming behaviour linked to low self-esteem. Finally, an abstract for CPR five is presented, summarising an oral case presentation describing an Acceptance and Commitment Therapy intervention for a 78-year-old man experiencing low mood following a stroke.

Acknowledgements

I would like to thank all of the children and families who kindly took part in the research. Particular thanks must go to the school staff who offered such enthusiasm, commitment and support to the recruitment process. Without your involvement, this research project would not have been possible.

I would also like to thank all those involved in supervising the project. Thank you to Hayley for your expert guidance and attention to detail, and for making the recruitment process so enjoyable, and thank you to Caroline for your support, encouragement and focus. I feel very lucky to have worked with such knowledgeable, inspiring and lovely people. Thank you to Chris O, for your expertise and feedback, and thank you to Chris J, for all your time, enthusiasm and statistics guidance.

Thank you to Alisa, for the dedication and effort you put into assisting with literature screening, your contribution has been invaluable.

And finally, thank you to my family, for the endless list of ways you have supported me throughout this journey. I am truly grateful.

Table of contents: Volume 1

Table of contents: Volume 2

Tables and Figures: Volume 1

CHAPTER 1

Tables

Figures

CHAPTER 2

Tables

Figures

Tables and Figures: Volume 2

CLINICAL PRACTICE REPORT 1

Figures

CLINICAL PRACTICE REPORT 2

Tables

Figures

CLINICAL PRACTICE REPORT 3

Tables

Figures

CLINICAL PRACTICE REPORT 4

Tables

Figures

CHAPTER 1

Exploring sources of variance in autism prevalence in rare syndromes: A meta-analysis

1. Abstract

Background: Elevated rates of autism spectrum disorder (ASD) have been reported in a range of rare genetic syndromes, prompting a focus on these syndromes for exploring genotype-phenotype links in ASD. A recent meta-analysis by Richards et al. (2015), however, observed significant variability in ASD prevalence estimates in 12 syndromes associated with elevated ASD phenomenology, suggesting factors other than syndrome diagnoses influenced prevalence estimates in these groups. The current analysis aimed to update and extend Richards et al.'s (2015) meta-analysis, evaluating the extent to which methodological factors and sample characteristics account for variance in ASD prevalence estimates.

Method: Richards et al.'s (2015) literature search was replicated in order to identify recent studies reporting ASD prevalence estimates in 21 syndromes associated with elevated ASD phenomenology. Pooled prevalence estimates were generated applying random-effects and quality-weighted models. Sub-group comparisons and meta-regression analyses evaluated the influence of individual methodological factors and sample characteristics on pooled prevalence estimates.

Results: The literature search generated 14 syndromes with sufficient data to produce robust pooled prevalence estimates. The odds of ASD classification were higher for all syndromes compared to general population estimates. After correcting for publication bias, ASD was highly prevalent (pooled prevalence estimates >30%) in tuberous sclerosis complex and Rett, Cohen, Angelman, Cornelia de Lange, Fragile X and CHARGE syndromes. Individual methodological factors had a significant influence on pooled prevalence estimates within individual syndrome groups but not when all syndrome data were combined. Across all syndrome data ASD prevalence was positively associated with

the proportion of the sample classified as having intellectual disability, and negatively associated with mean age and IQ.

Discussion: The contribution of the results for delineating the relationship between genetic syndromes and ASD phenomenology is discussed. Implications for clinical practice and future research are also explored.

2. Introduction

Autism spectrum disorder (ASD) describes a developmental disorder estimated to occur in approximately 1% of the general population (Baird et al., 2006; Baron-Cohen et al., 2009). ASD is characterised by communication difficulties, impairments in social interaction, and the presence of restricted and repetitive interests and behaviours (American Psychiatric Association, 2013; World Health Organisation, 1992).

ASD is a behaviourally classified disorder, of which the genetic aetiology is unclear. Numerous genetic loci, de-novo mutations and copy-number variations have been associated with an increased risk of ASD (Talkowski, Minikel, & Gusella, 2014), indicating that diagnostic behavioural characteristics likely arise from interactions between multiple genetic and environmental risk factors (Persico & Bourgeron, 2006; Zhao et al., 2007). In an attempt to delineate these complex mechanisms, increasing attention has been paid to known genetic syndromes associated with elevated ASD symptomatology (Betancur & Buxbaum, 2013; Constantino et al., 2015; Garg et al., 2015). Examples include Fragile X syndrome, Rett syndrome and tuberous sclerosis complex, all of which evidence higher ASD prevalence rates than is expected in the general population (Moss & Howlin, 2009; Oliver, Berg, Moss, Arron, & Burbidge, 2011). Lee, Martin, Berry-Kravis and Losh (2016) suggest that genetic syndromes offer a "simplified context" within which to explore genotype-phenotype links, with the potential for developing new models for understanding the biological pathways underpinning nonsyndromic ASD.

Exploring genotype-phenotype links in syndromic ASD necessitates a well-defined account of ASD phenomenology in each syndrome; however, reported ASD prevalence

estimates vary considerably, both within and across syndrome groups (Richards, Jones, Groves, Moss, & Oliver, 2015). Furthermore, emerging evidence suggests that the specific behavioural profile in individuals with rare genetic syndromes might differ from that observed in nonsyndromic ASD. Individuals with Rett syndrome, for example, have been observed to show preserved eye contact, despite impairments in other areas of social interaction (Nomura & Segawa, 2005). In contrast, Cornish, Turk and Levitas (2007) suggest that individuals with Fragile X syndrome are socially motivated, with a good understanding of social interaction, but display gaze avoidant behaviour due to social anxiety and hypersensitivity. These findings indicate a need for more precise and detailed delineation of ASD phenomenology in these genetic syndromes.

Noting the lack of synthesised data regarding ASD phenomenology in rare genetic syndromes, Richards et al. (2015) sought to generate robust ASD prevalence estimates in 21 genetic syndromes associated with ASD (based on Moss & Howlin's, 2009, review). Twelve syndromes were identified as having a sufficient number of papers for review. Following the application of quality criteria, quality-weighted pooled prevalence estimates of ASD symptomatology were elevated for all 12 syndrome groups compared to general population estimates. Richards et al. (2015) reported that Rett syndrome, Cohen syndrome, tuberous sclerosis complex, Cornelia de Lange syndrome, Angelman syndrome and CHARGE syndrome were associated with a “high” prevalence of ASD. As such, these syndromes may provide a reliable focus for research exploring gene-behaviour links in ASD (Richards et al., 2015). Richards et al. (2015) further suggested that their findings highlighted a need for clinical and educational services to provide comprehensive assessment and specific support for ASD-linked difficulties in these syndrome groups.

However, an important issue not addressed by Richards et al.'s (2015) analysis was the extent to which participant characteristics beyond syndrome-specific factors might influence prevalence estimates. Indeed, the prevalence and severity of ASD has been shown consistently to correlate negatively with intellectual disability, prompting some to question whether intellectual disability present within genetic syndromes, rather than syndrome-specific characteristics per se, underpins the link between genetic syndromes and ASD phenomenology (Moss & Howlin, 2009). Skuse (2007) suggests that intellectual disability might limit an individual's capacity to compensate for autistic vulnerabilities, and therefore acts as an additional risk factor in the development of ASD-linked behaviour. In support of this argument, studies have reported negative associations between IQ and ASD severity in individuals with Fragile X syndrome (Hall, Lightbody, Hirt, Rezvani, & Reiss, 2010; Loesch et al., 2007; Thurman, McDuffie, Kover, Hagerman, & Abbeduto, 2015) and Angelman syndrome (Trillingsgaard & Østergaard, 2004). Similarly, Molloy et al. (2009) found intellectual ability was lower in individuals with Down syndrome and ASD, compared to individuals with Down syndrome only. In contrast, however, Moss and Howlin's (2009) review of research exploring the prevalence of ASD in tuberous sclerosis complex concluded that intellectual disability alone could not account for elevated ASD prevalence, with prevalence estimates of up to 17% in individuals with tuberous sclerosis complex who have IQ scores in the typical range (de Vries, Hunt, & Bolton, 2007; Prather & de Vries, 2004). The extent to which intellectual disability influences ASD prevalence estimates across and within syndrome groups is therefore unclear. Further exploration of this association is required in order to understand the relationship between genetic syndrome diagnoses and ASD phenomenology.

In addition to the intellectual ability of study participants, the proportion of males to females in participant samples might also be hypothesised to influence prevalence estimates. The prevalence of nonsyndromic ASD is consistently reported to be higher in males than females, with a generally accepted male to female ratio of 4:1 (Baird et al., 2006). As a result, participant samples with a high male to female ratio might be expected to show elevated rates of ASD, regardless of syndrome-specific factors. This has important implications when studying genetic syndromes with significantly skewed gender ratios. Fragile X syndrome is up to two times more common in males than females (Crawford, Acuña, & Sherman, 2001; Tassone et al., 2012), whilst almost all individuals diagnosed with Rett syndrome are female (Wulffaert, Van Berckelaer-Onnes, & Scholte, 2009). The extent to which these gender differences accounted for heterogeneity in reported prevalence rates was not explored in Richards et al.'s (2015) meta-analysis.

Age-related differences have also been noted in nonsyndromic ASD, with studies suggesting a decline in ASD severity through adolescence and adulthood (Seltzer, Shattuck, Abbeduto, & Greenberg, 2004; Woodman, Smith, Greenberg, & Mailick, 2015). Woodman et al. (2015) reported an overall reduction in ASD phenomenology in adolescents and adults with idiopathic ASD over an 8.5 year period, with changes particularly evident in the domains of verbal communication and restricted and repetitive behaviours. Studies have also reported age-related changes in ASD symptomology in some syndrome groups. In Rett syndrome there is some evidence of an improvement in ASD symptoms over time (Nomura & Segawa, 2005). In contrast, Hatton et al. (2006) and Lee et al. (2016) both report positive associations between age and severity of ASD characteristics in children and adolescents with Fragile X syndrome. The mean age of participant samples in Richards et al.'s (2015) meta-analysis varied substantially between

studies, thus age may have been an important moderating variable. The effect of participant age on ASD prevalence estimates in genetic syndromes therefore needs further exploration.

In addition to demographic characteristics, it is also plausible that measurement characteristics influence ASD prevalence estimates. Evidence suggests that ASD prevalence estimates are higher when individuals are assessed against the DSM-IV-TR diagnostic criteria (American Psychiatric Association, 2000) compared to the DSM-5 (American Psychiatric Association, 2013), for example (Hartley et al., 2015). Furthermore, diagnostic tools may provide less specificity when administered with children with intellectual disability (Gray, Tonge, & Sweeney, 2008). Again, Richards et al. (2015) noted significant variability in ASD assessment, ranging from parental report or screening tools to comprehensive assessment using multiple 'gold-standard' diagnostic measures. Thus, there is a need to evaluate the influence of this variability on ASD prevalence estimates.

The current study therefore had two key aims:

1) To update the pooled prevalence estimates generated by Richards et al. (2015), incorporating the wealth of research published in the three years since their original analysis.
2) To evaluate sources of variance in ASD prevalence estimates, both within and between syndrome groups. This included examining the influence of methodological factors, for example ASD assessment method, and sample characteristics, such as age, intellectual ability and gender.

3. Method

3.1 Search strategy

The search results generated by Richards et al. (2015) were updated through a literature search of the PyschINFO, Medline, Embase and Pubmed Central computerised databases. Replicating Richards et al.'s (2015) search strategy, the search included 21 genetic syndromes, identified as being associated with ASD phenomenology in a review by Moss and Howlin (2009; see Richards et al., 2015, for further information). To ensure that no papers were overlooked, search dates overlapped with those of Richards et al. (2015) by a minimum of one month. Details of the syndromes investigated, search terms and search dates are provided in Appendix 1. The search identified all articles with any of the syndrome search terms for each syndrome group 'AND' any of the following ASD search terms: Autis*, Autism*, Autism Spectrum Disorder*, ASD, Pervasive Developmental Disorder*, Pervasive Developmental Disorder not otherwise specified, PDD-NOS, PDDNOS, Unspecified PDD, Asperger*, and Asperger* syndrome.

3.2 Selection strategy

Papers identified through the literature search were screened and selected according to the following selection strategy. Syndrome groups were removed if, at any point in the selection process, the number of papers available for that group, in combination with the existing papers from Richards et al.'s (2015) meta-analysis, was fewer than two. A summary of the number of papers included and excluded at each phase of the selection process is provided in Figure 1, following the PRISMA guidelines for conducting and reporting meta-analyses (Moher, Liberati, Tetzlaff, Altman, & The PRISMA Group, 2009).

Further detail regarding the selection strategy and reasons for exclusion are provided in Appendix 2.

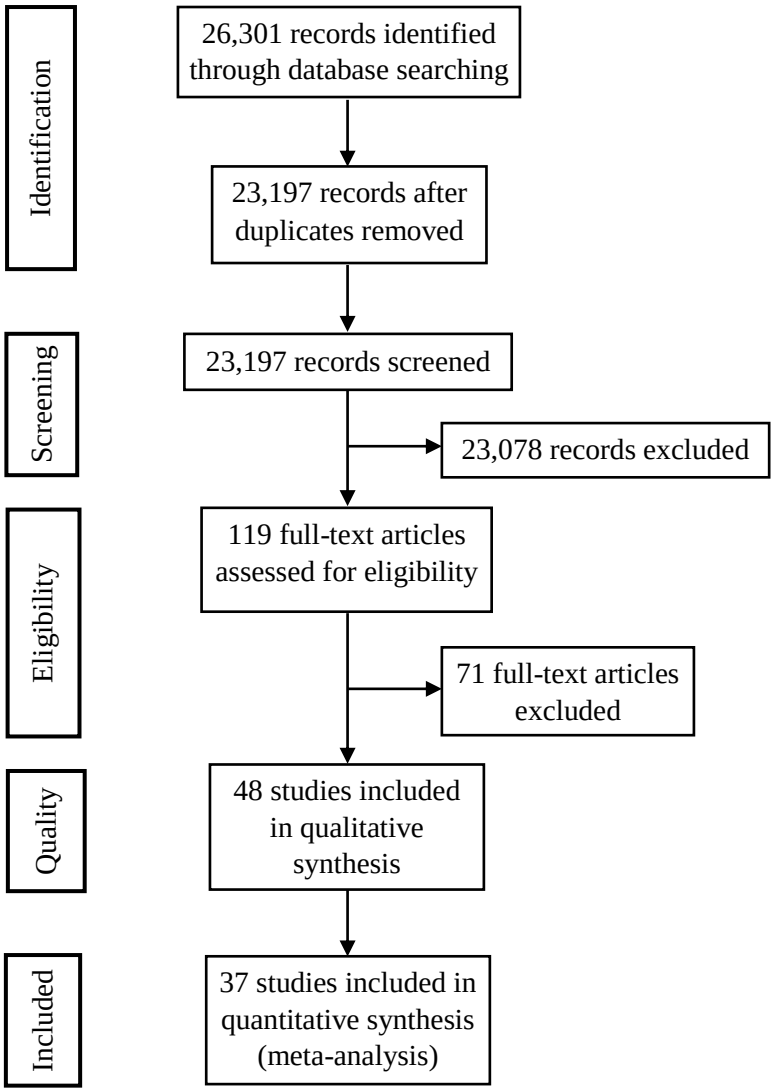

Figure 1. Selection of studies for review

3.2.1 Screening

The title and abstract of all identified papers were screened by a single researcher, applying the inclusion and exclusion criteria detailed in Table 1. Where the first researcher was

uncertain regarding a paper's inclusion/exclusion, the paper was reviewed and discussed with a second researcher until consensus was reached.

Table 1. Inclusion and exclusion criteria applied at the screening stage

Inclusion criteria	Exclusion criteria
Empirical papers published in a peer reviewed journal	Dissertations, conference proceedings, magazine articles, review articles/discussion papers and books
Papers available in English	Papers not available in English
The title/abstract states that the paper reports on ASD diagnosis/symptomatology in the named syndrome group	Sample recruited based on existing/ suspected ASD diagnosis or symptomatology
Sample size of the syndrome group ≥10	Sample size of the syndrome group ≤10

3.2.2 Eligibility

All articles identified as suitable at the screening stage were read in full by a single researcher. Eligibility for inclusion in the analysis was determined according to the inclusion/criteria provided in Table 1, along with the additional criteria specified in Table 2 below. Again, where eligibility could not be determined by the first researcher, a second researcher was asked to review the paper and consensus was achieved.

Table 2. Additional inclusion and exclusion criteria applied at the eligibility stage

Inclusion criteria	Exclusion criteria
The study reports the number/percentage of individuals who meet specified criteria for ASD	The study only reports whole group averages on a measure of ASD symptomatology
No apparent bias in participant selection	Participants were selected based on the presence/absence of a particular characteristic, e.g. seizures/epilepsy, intellectual ability, sensory difficulty/disorder[1]
The study reports on an original sample, or the degree of overlap with another sample could not be established	The study reports on the same sample as, or a sub-sample of, a previous study

[1] In syndromes with genetic mechanisms associated with an unequal gender distribution (e.g. Rett syndrome, Fragile X syndrome) studies with samples selected on the basis of gender were not excluded.

3.2.3 Quality

Papers considered eligible were evaluated according to the quality criteria produced by Richards et al. (2015). These criteria were developed through a review of existing literature and quality criteria, in addition to consultation with relevant research experts (see Richards et al., 2015, for full details of the development process). According to this criteria, studies were assigned individual numerical ratings between 0 and 3 for each of the following categories: a) Sample Identification (i.e. how individuals with the syndrome were selected), b) Confirmation of Syndrome (i.e. the method though which participants were confirmed to belong to the syndrome group), and c) ASD Assessment (i.e. the method/measures through which ASD diagnosis was determined). The criteria for each quality rating are provided in Table 3.

Based on the ratings for each of the three quality criteria, papers were further assigned an overall quality rating, generated by dividing the total quality rating by nine (the maximum total score). Papers receiving an overall quality rating ≥0.33, attained across at least two quality criteria, were included in the analysis. A sample of 15 (31.15%) papers were reviewed and rated by a second researcher. Inter-rater reliability was good for Sample Identification ($r_s(13)=.945$, $p=<.001$), Confirmation of Syndrome ($r_s(13)=.706$, $p=.003$), ASD Assessment ($r_s(13)=.756$, $p=.001$) and overall quality ratings ($r_s(13)=.767$, $p=.001$).

Table 3. Quality ratings for Sample Identification, Confirmation of Syndrome and ASD Assessment quality criteria

	Quality Rating			
	0 **Poor**	**1** **Adequate**	**2** **Good**	**3** **Excellent**
Sample Identification	Not specified/reported OR Parent/carer/family/self-report only	Single restricted or non-random sample, e.g. a specialist clinic or previous research study[1] Single regional sample, e.g. a regional parent support group	Multiple restricted or non-random samples, e.g. multi-region specialist clinics National non-random sampling, e.g. national parent support groups	Random or total population sample
Confirmation of Syndrome	Not confirmed/reported OR parent/carer/family/self-report only Clinical diagnosis only suspected	Clinical diagnosis by 'generalist', e.g. General Practitioner or Paediatrician	Clinical diagnosis by 'expert', e.g. Clinical Geneticist or Specialist Paediatrician	Molecular/Cytogenetic/ Metabolic confirmation of diagnosis[2]
ASD Assessment	Not specified/reported Clinician judgement only	Screening instrument, e.g. SCQ, M-CHAT Clinician judgement against specified diagnostic criteria, e.g. DSM-IV or ICD-10	Diagnostic instrument, e.g. ADI-R, DISCO, ADOS, 3Di	Consensus from multiple assessments, including at least one diagnostic instrument

[1]For individuals recruited as part of a larger *ongoing* study, if the recruitment strategy is described, it is coded. If not, it is coded as 1, indicating the sample has come from one source (i.e., the larger ongoing study).

[2] For syndromes where genetic causes are only currently identified for a proportion of cases (e.g., in CdLS, the NIP-BL gene deletion is thought to account for only 50% of cases), the study will receive a score of 3 if they tested all participants, even if all participants did not evidence the genetic marker and were subsequently confirmed through clinical assessment of features.

3.3 Data analysis

All analyses were conducted using Metafor and Meta programs with the R analysis package (Schwarzer, 2007; Viechtbauer, 2010).

3.3.1 Pooled prevalence estimates

The first aim of the study was to update the pooled prevalence estimates reported in Richards et al. (2015). Prevalence data (i.e. the number of individuals meeting ASD criteria) were therefore extracted from each paper, and added to Richards et al.'s (2015) existing data. In line with Richards et al.'s (2015) analysis, prevalence values were extracted according to procedure outlined in Figure 2.

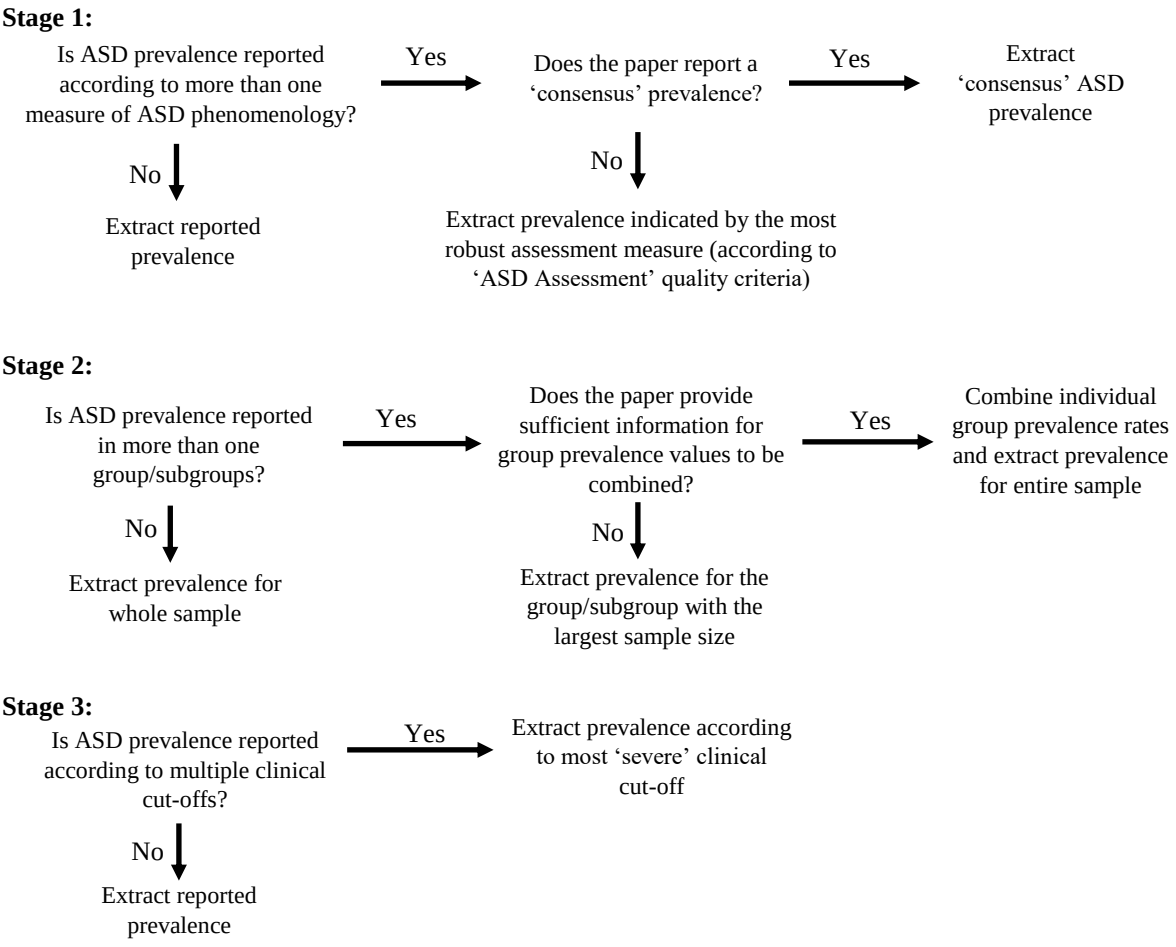

Figure 2. Procedure for extraction of ASD prevalence data

Pooled prevalence estimates were then generated using a random-effects model. Original, fixed-effects models of meta-analysis assume that variability between studies is a result of sampling error only, and that a true, common outcome exists across all studies (Lipsey & Wilson, 2001; Senn, 2007). In contrast, random-effects models assume that study level differences produce a further source of variability, and weights the contribution of studies accordingly (Barendregt & Doi, 2011). Given the variability in prevalence rates reported between studies, including within syndrome groups, a random-effects model was deemed most appropriate.

Despite the advantages of the random-effects model, it neglects to consider variability between studies as a result of methodological differences other than sample size (Barendregt & Doi, 2011). A quality-effects model of analysis, which accounts for these sources of bias by weighting the contribution of each study according to explicit ratings of methodological quality (Barendregt & Doi, 2011), was therefore also applied. Studies were weighted according to the overall numerical quality rating generated at the Quality stage of the selection strategy, using Richards et al.'s (2015) quality criteria.

In order to evaluate the relative risk of ASD between syndrome groups, updated relative risk statistics were generated based on pooled prevalence estimates. Odds ratios were further calculated to compare the odds of ASD in each syndrome group compared to general population estimates.

3.3.2 Examining sources of variance

Syndrome groups were examined in turn, to explore sources of heterogeneity in the data. Two methods were used to identify individual studies exerting an influence on both the

overall meta-analytic result and heterogeneity. These were a) visual inspection of a Baujat plot (Baujat, Mahé, Pignon, & Hill, 2002), and b) a leave-one-out analysis, in which each study was excluded in turn and the impact of its omission on the meta-analytic result and data heterogeneity was evaluated.

In order to evaluate the influence of methodological factors, sub-group[1] analyses using the random-effects model assessed the effect of quality ratings for individual quality criterion on ASD prevalence estimates, both within individual syndromes and across all data. A series of univariate meta-regression analyses were also conducted to evaluate the influence of sample characteristics on prevalence estimates, including the proportion of male participants in the sample, the mean age of the sample, and degree of intellectual disability (assessed both by the proportion of the sample with an intellectual disability and the mean IQ of the sample). Again, meta-regression analyses were calculated for individual syndromes and across data from all syndrome groups.

3.3.3 Differences between the current analysis and Richards et al. (2015)

The meta-analysis by Richards et al. (2015) was carried out using MetaXL (Barendregt & Doi, 2011), whilst the current analysis used Metafor and Meta programs using the R analysis package (Schwarzer, 2007; Viechtbauer, 2010). These libraries are open source and peer reviewed, and are considered to be valid methods of meta-analysis.

The current study applied a log odds transformation to normalise the distribution of study effects, which was not applied in the original analysis by Richards et al. (2015). As a result

[1] 'Sub-group' analyses refer to analyses in which separate pooled prevalence estimates were generated for studies grouped according to a particular variable, for example the quality rating assigned for a single quality criterion, and the difference between these pooled prevalence estimates was assessed for statistical significance.

of these changes, there may be very small differences between the estimates of prevalence from the original and current analysis.

4. Results

4.1 Overview

Studies generated by the literature search were first reviewed, and their contribution was evaluated with reference to methodological quality. Updated pooled prevalence estimates were then generated, combining new studies with the existing research identified by Richards et al. (2015). Updated relative risk statistics and odds ratios, comparing ASD prevalence between syndromes and with the general population, were also calculated.

In order to explore sources of variance in ASD prevalence estimates, the impact of influential studies, individual quality criterion and sample characteristics on prevalence estimates were examined for each syndrome group in turn. Analyses exploring the impact of quality criteria and sample characteristics on the variance in ASD prevalence estimates across *all* syndromes were also undertaken.

4.2 Identified research

The literature search identified a total of 48[1] studies describing ASD prevalence in genetic syndromes. These studies are discussed below and summarised in Tables 4-15. No new studies were found reporting the prevalence of ASD in individuals with Cohen, Joubert, Moebius or Rett syndromes. The literature searches for studies reporting ASD prevalence estimates in CHARGE syndrome and Williams syndrome each identified one new study

[1] One further study was identified reporting ASD prevalence in Ehlers-Danlos syndrome, however, this was excluded as fewer than two papers were available for this group.

(see Tables 4 and 5), however, these studies did not meet the pre-set quality criteria and were therefore excluded at the Quality stage of the review.

Table 4. Quality criteria, participants and study characteristics of research reporting ASD prevalence estimates in CHARGE syndrome

	Quality Criteria		*CHARGE Syndrome Study and Sample Characteristics*										*Outcome Data*	
Authors	**Sample**	**Syndrome ASD**	**N**	**% Male**	**Mean Age *(SD)* *Range***	**% with ID**	**Mean IQ**	**Syndrome Diagnosis**	**ASD Measure**	**ASD Measure Professional**	**ASD Profile**	**%ASD categories**	**% ASD (*N*)**	**Quality Weighting**
Hartshorne *et al.*, 2016			53	62	*13-39*	85[1]	Not reported	Self-report	Self-report	Participants and/or caretakers	No	Autism spectrum: 26	26 *(14)*	0.22

[1]Developmental delay

Table 5. Quality criteria, participants and study characteristics of research reporting ASD prevalence estimates in Williams syndrome

	Quality Criteria		*Williams Syndrome Study and Sample Characteristics*										*Outcome Data*	
Authors	**Sample**	**Syndrome ASD**	**N**	**% Male**	**Mean Age *(SD)* *Range***	**% with ID**	**Mean IQ**	**Syndrome Diagnosis**	**ASD Measure**	**ASD Measure Professional**	**ASD Profile**	**%ASD categories**	**% ASD (*N*)**	**Quality Weighting**
Reilly *et al.*, 2015			80	50	11.06 *(3.96)* *14.7*	Not reported	Not reported	Not reported	Parental report of professional diagnosis	N/A	No	ASD (males): 5 ASD (females): 3 ASD (whole sample): 4	4 *(3)*	0.22

4.2.1 Fragile X syndrome

Thirteen new studies were identified, however, four did not meet pre-set quality criteria and were therefore excluded (see Table 6). Prevalence estimates were varied, ranging from 20-59%. Overall, quality ratings were good. Nine studies achieved an Excellent quality rating for Confirmation of Syndrome, however, the quality of ASD Assessment was mixed (ranging Adequate to Excellent), and no studies were rated as Excellent for Sample Identification. Studies by Kaufmann et al. (2017) and Wheeler et al. (2015) are notable for their large sample sizes, however, both achieved an Adequate rating only for ASD Assessment, relying on clinician judgement against DSM-IV/DSM-5 diagnostic criteria.

Table 6. Quality criteria, participants and study characteristics of research reporting ASD prevalence estimates in Fragile X syndrome

	Quality Criteria	*Fragile X Syndrome Study and Sample Characteristics*										*Outcome Data*	
Authors	**Sample Syndrome ASD**	**N**	**% Male**	**Mean Age *(SD)* *Range***	**% with ID**	**Mean IQ**	**Syndrome Diagnosis**	**ASD Measure**	**ASD Measure Professional**	**ASD Profile**	**%ASD categories**	**% ASD (*N*)**	**Quality Weighting**
Adlof *et al.*, 2015		54	100	10.26 *(1.67)* *7.67-14.04*	Not reported	5.31[1]	DNA analysis[2]	CARS[3]	Not reported	No	Autism:35	35 *(19)*	0.22
Grefer *et al.*, 2016		33	100	T1: 41.09[4] *(6.50)* *36-59* T2: 64.79 *4.46* *60-83*	Not reported	Not reported	Genetic	CARS	Examiner	No	ASD (T1): 36 ASD (T2): 33	33 *(11)*[5]	0.67
Greiss Hess *et al.*, 2016		57	84.2	Group 1:[6] 3.89 *(1.09)* Group 2: 3.92 *(1.11)*	Not reported	Group 1:[7] 56.6 Group 2: 54.8	Genetic	ADOS-2 DSM-5	Not reported	No	ASD: 59.2	59.2 *(32)*	0.67
Hall *et al.*, 2017		148	100	11.56 *(2.6)* *8-16*	Not reported	Not reported	Parental report	Parental report	N/A	No	ASD: 54.7	54.7 *(81)*	0.22

[1] Mean non-verbal mental age based on the Brief IQ composite score from the Leiter International Performance Scales-Revised
[2] DNA analysis available for 42 participants. No information regarding syndrome confirmation for entire sample
[3] Childhood Autism Rating Scale. Data available for 50/54 participants
[4] Age in months, sample assessed at two time points
[5] ASD prevalence at Time 2 was extracted as this provided the more conservative estimate
[6] Participants randomly assigned to two groups: Sertraline (Group 1) and placebo (Group 2)
[7] Mean Mullen Scales of Early Learning Composite score

Authors	*Quality Criteria* Sample	Syndrome	ASD	N	% Male	Mean Age *(SD)* *Range*	% with ID	Mean IQ	Syndrome Diagnosis	ASD Measure	ASD Measure Professional	ASD Profile	%ASD categories	% ASD (*N*)	Quality Weighting
	Fragile X Syndrome Study and Sample Characteristics													*Outcome Data*	
Kaufmann *et al.*, 2017				564	78	*3-21*	Not reported	Not reported	Genetic	DSM-IV/DSM-5	Clinic physician	No	ASD: 42	42 *(237)*	0.67
Klusek *et al.*, 2015				51	100	10.2 *7.9-13.2*	Not reported	56.0[8]	DNA analysis	CARS	Trained research associates	No	ASD: 20	20 *(10)*	0.67
Lisik *et al.*, 2015				23	100	19.87 *(6.56)*	Mild: 4.3 Moderate: 47.8 Profound: 47.8	Not reported	DNA analysis	DSM-IV	Not reported	No	Autism: 21.7	21.7 *(5)*	0.56
Pretto *et al.*, 2014				18	77.8	32.2 *13-73*	≤70: 61.1	66.6	Genetic	ADOS	Not reported	No	ASD: 50	50 *(9)*	0.67
Reilly *et al.*, 2015				115	82	11.58 *(3.6)* *15.26*	Not reported	Not reported	Not reported	Parental report of professional diagnosis	N/A	No	ASD (males): 44 ASD (females): 14 ASD (whole sample): 39	39 *(43)*	0.22
Roberts *et al.*, 2016				15	100	12.33 *(1.27)* *9.37-14.35*	Not reported	69.8[9]	Genetic report	AOSI[10] ADOS toddler[11]	Research staff trained to research reliability standards	Yes	AOSI autism risk threshold: 53 ADOS: 40	40 *(4)*	0.67

[8] Mean Brief IQ standard score on the Leiter International Performance Scale-Revised
[9] Mean Mullen Scales of Early Learning Standard Score
[10] Autism Observation Scale for Infants – identifies signs of autism in infants aged 6-18 months
[11] Completed for 10 participants who had reached 24 months of age

Authors	*Quality Criteria* Sample	Syndrome	ASD	N	% Male	Mean Age *(SD)* *Range*	% with ID	Mean IQ	Syndrome Diagnosis	ASD Measure	ASD Measure Professional	ASD Profile	%ASD categories	% ASD (*N*)	Quality Weighting
	Fragile X Syndrome Study and Sample Characteristics													*Outcome Data*	
Russo-Ponsaran *et al.*, 2014				11	0	11.23 *(2.92)* *7.2-18.0*	≤70: 36.4	76.8[12]	DNA analysis	SRS ADOS	Doctoral/ bachelor's level research staff Research reliable doctoral level researcher	No	Consensus autism spectrum: 45.5	45.5 *(5)*	0.78
Warren *et al.*, 2017				55	80	34.11[13] *(5.58)* *24-55*	Not reported	44.31[14]	Not reported	Maternal report	Paediatrician, paediatric neurologist, psychologist	No	Autism: 32.7	32.7 *(18)*	0.22
Wheeler *et al.*, 2015				758	84.3	Males 19.6 *2-67* Females 14.6 *3-44*	Not reported	Not reported	Caregiver report	Caregiver reported diagnosis Autism survey, DSM-IV-TR/ DMS-5[14]	Caregiver report Expert diagnostician, confirmed by second clinician-researcher	Yes	Autistic disorder: 22 DSM-IV-TR AD/ASD: 38.7(m), 24.7(f) DSM-5[15] AD/ASD: 27.8(m), 11.3(f)	27.8 *(178)*	0.33

[12] Wechsler Abbreviated Scale of Intelligence

[13] Age in months

[14] Mean Mullen Scales of Early Learning nonverbal raw score

[14] Current behaviours, as assessed by an 'autism survey' completed by parents, were mapped onto DMS-IV-TR/DSM-5 criteria by an 'expert diagnostician'

[15] Prevalence estimates were extracted for the male group, assessed against DSM-5 criteria; this provided a prevalence estimate for the largest participant sample and according to the most recent revision of the diagnostic criteria.

4.2.2 22q11.2 deletion syndrome

Nine studies reported the prevalence of ASD in 22q11.2 deletion syndrome, although one study was removed due to low quality ratings (see Table 7). Of note, five studies reported prevalence estimates ≥40%, which is significantly higher than the pooled prevalence estimate of 11% reported by Richards et al. (2015). The remaining studies were consistent with this earlier estimate. Overall, methodological quality ratings were very good. All eight studies were assigned an Excellent rating for Confirmation of Syndrome. Furthermore, four studies also used multiple methods of ASD assessment, and classified ASD based on consensus between these measures.

Table 7. Quality criteria, participants and study characteristics of research reporting ASD prevalence estimates in 22q11.2 deletion syndrome

	Quality Criteria		*22q11.2 Deletion Syndrome Study and Sample Characteristics*										*Outcome Data*	
Authors	**Sample**	**Syndrome ASD**	**N**	**% Male**	**Mean Age *(SD)* *Range***	**% with ID**	**Mean IQ**	**Syndrome Diagnosis**	**ASD Measure**	**ASD Measure Professional**	**ASD Profile**	**%ASD categories**	**% ASD (*N*)**	**Quality Weighting**
de Sonneville *et al.*, 2016			58	34.5	13.5 *(2.6)* *9.0 – 18.5*	Not reported	65.2	Genetic	ADI-R DSM-IV	MDT consensus meeting, led by child psychiatrist	Yes	ASD: 53.4	53.4 *(31)*	0.89
Fiksinski *et al.*, 2017			89	40.4	14.3 *(1.9)* *11.3-18.9*	Not reported	64.1[1]	Genetic	DSM-IV ADI-R	MDT Certified interviewers	No	ASD: 58.4 *Subgroups* Autistic disorder: 5.6 PDD-NOS: 52.8	5.6 *(5)*	0.78
Fjermestad *et al.*, 2015			12	25	14.5 *(1.4)* *12.0-17.0*	Not reported	Not reported	Genetic	Kiddie-SADS screening	Clinical Psychologists	No	ASD: 50	50 *(6)*	0.56
Hidding *et al.*, 2015			102	39.2	13.2 *(2.6)* *9-18.5*	Not reported	66.0	Genetic	DSM-IV ADI-R	MDT consensus meeting, headed by child psychiatrist Certified interviewers	Yes	ASD: 48.0	48.0 *(49)*	0.89

[1] IQ data available for 80 participants

Authors	*Quality Criteria* Sample	Syndrome	ASD	N	% Male	Mean Age *(SD)* *Range*	% with ID	Mean IQ	Syndrome Diagnosis	ASD Measure	ASD Measure Professional	ASD Profile	%ASD categories	% ASD (*N*)	Quality Weighting
	22q11.2 Deletion Syndrome Study and Sample Characteristics													*Outcome Data*	
Hidding *et al.*, 2016				45	40	13.3 *(2.7)* *9-18.5*	Not reported	66.3	Genetic	DMS-IV ADI-R	MDT consensus meeting, headed by child psychiatrist Certified interviewers	Yes	ASD: 55.6	55.6 *(25)*	0.89
Jalbrzikowski *et al.*, 2015				40	50	17.3 *(11.9)*	Not reported	76.6	Genetic	≤18 years: ADOS, ADI-R >18 years: SCID+ developmental disorders module[2]	Master's- and PhD-level clinicians	No	ASD: 40	40 *(16)*	0.67
Olszewski *et al.*, 2017				57	54.4	20.87 *(2.29)*	Not reported	74.54	Cytogenetic	SRS	Not reported	No	ASD: 12.3	12.3 *(7)*	0.56
Reilly *et al.*, 2015				76	56	10.39 *(3.41)* *13.66*	Not reported	Not reported	Not confirmed	Parental report of professional diagnosis	N/A	No	ASD (males): 7 ASD (females): 6 ASD (whole sample): 6.6	6.6 *(5)*	0.22
Vangkilde *et al.*, 2016				29	65.5	15.7 *(2.8)*	17.2	79.52	Molecular	SCQ	Not reported	No	ASD: 13.8	13.8 *(4)*	0.67

[2] SCID: Structured Clinical Interview for DSM-IV Axis I disorders

4.2.3 Tuberous sclerosis complex

The literature search identified nine new studies reporting the prevalence of ASD in tuberous sclerosis complex, however, two studies, by Taquet et al. (2014) and Tye et al. (2015), did not meet the required quality rating and were excluded (see Table 8). Again, substantial variance was observed in prevalence estimates, which ranged from 17-55%. Overall, Confirmation of Syndrome ratings were low; three studies relied on clinical diagnosis by a 'generalist' and one study did not report syndrome confirmation. However, two studies (Huang et al., 2015, and Yang et al., 2017) included genetic criteria. No studies were rated as Excellent for either Sample Identification or ASD Assessment. The study by Kothare et al. (2014) reported a large sample size (*n*=916), with data obtained from a multi-centre database. However, the method of ASD assessment was not reported.

Table 8. Quality criteria, participants and study characteristics of research reporting ASD prevalence estimates in tuberous sclerosis complex

	Quality Criteria		*Tuberous sclerosis complex Study and Sample Characteristics*									*Outcome Data*	
Authors	**Sample Syndrome ASD**	**N**	**% Male**	**Mean Age *(SD) Range***	**% with ID**	**Mean IQ**	**Syndrome Diagnosis**	**ASD Measure**	**ASD Measure Professional**	**ASD Profile**	**%ASD categories**	**% ASD (*N*)**	**Quality Weighting**
Huang *et al.*, 2015		32	50	*0.08-45*	43.8[1]	Not reported	Genetic	Not reported	Not reported	No	ASD: 18.8	18.8 *(6)*	0.44
Jeste *et al.*, 2014		40	Not reported	Not reported[2]	Not reported	Not reported	Not reported	ADOS	Trained research assistants	No	ASD: 55	55 *(22)*	0.44
Jeste *et al.* 2016		44	62	32.1[3] *23-39*	Not reported	Not reported	Clinical presentation[4]	ADOS	Not reported	No	ASD: 50[5]	50 *(18)*	0.56
Kothare *et al.*, 2014		916	49	50[6] *(90)*	18.7	Not reported	Clinical[7]	Not reported[8]	Not reported	No	ASD: 16.9	16.9 *(155)*	0.33

[1] Data available for 17 participants
[2] Study reports longitudinal data, with ADOS assessments conducted at multiple time-points. The last ADOS score was used for ADOS classification.
[3] Age in months
[4] The study did not indicate whether diagnosis was determined by a generalist or specialist, therefore a conservative quality rating of one was applied for 'Syndrome Confirmation'.
[5] Full data available for 36 participants only
[6] Age in months
[7] Data from a patient database. Unknown whether diagnoses were determined by a generalist or specialist, therefore a conservative quality rating of one was applied for 'Syndrome Confirmation'.
[8] Not specified; clinical data from a patient registry

Authors	*Quality Criteria* Sample	Syndrome	ASD	N	% Male	Mean Age *(SD)* *Range*	% with ID	Mean IQ	Syndrome Diagnosis	ASD Measure	ASD Measure Professional	ASD Profile	%ASD categories	% ASD (*N*)	Quality Weighting
	Tuberous sclerosis complex Study and Sample Characteristics													*Outcome Data*	
Shehata *et al.*, 2017				36	52.8	6.5 *(2.1)* *2.5-10.9*	75	Not reported	Clinical	ADOS	Not reported	No	ASD: 30.6	30.6 *(11)*	0.56
Taquet *et al.*, 2014				38	Not reported	Not reported	Not reported	Not reported	Not reported	Not reported	Not reported	No	ASD: 26.3[9]	26.3 *(10)*	0.00
Tye *et al.* 2015				14	64.3	25.5[10] *(8.52)* 22.33 *(6.92)*	No reported	88.29[11] 53.33	Not reported	Clinical assessment, DSM-IV	MDT	No	ASD: 42.9	42.9 *(6)*	0.11
Wilbur *et al.*, 2017				81	51	10.0[12] *0.2-23.2*	33	Not reported	Clinical	Clinical document	Paediatrician, psychiatrist or neurologist	No	ASD: 25	25 *(20)*	0.33
Yang *et al.*, 2017				117	51.3	5.17 ±3.6	71.8	Not reported	Genetic	ABC[13]	Parental/ caretaker report	No	Autism: 23.1	23.1 *(27)*	0.44

[9] 11 participants "too young for diagnosis"
[10] Age reported separately for participants with TSC only (top) and TSC+ASD (bottom)
[11] Mean IQ for participants with TSC only (top) and TSC+ASD (bottom)
[12] Median age at most recent clinical follow-up
[13] Autism Behaviour Checklist

4.2.4 Down syndrome

The literature search identified four new studies providing ASD prevalence estimates in individuals with Down syndrome (see Table 9). The studies by Hoffmire, Magyar, Connolly, Fernandez and van Wijngaarden (2014) and Oxelgren et al. (2017) are notable for their high prevalence estimates compared to the pooled prevalence estimates generated by Richards et al.'s (2015) analysis. Both these studies received good overall quality ratings (0.56). The study by Oxelgren et al. (2017) was the only study to classify ASD based on consensus between multiple assessment measures, including both the ADOS (Gotham, Risi, Pickles, & Lord, 2007) and ADI-R (Lord, Rutter, & Le Couteur, 1994). Studies by Naerland, Bakke, Storvik, Warner and Howlin (2017) and Warner, Moss, Smith and Howlin (2014) both reported large sample sizes, although these studies had low quality ratings for Syndrome Identification and ASD Assessment; neither study reported genetic confirmation of Down syndrome, and both studies used a screening tool only (the Social Communication Questionnaire; Rutter, Bailey, & Lord, 2003) for ASD classification.

Table 9. Quality criteria, participants and study characteristics of research reporting ASD prevalence estimates in Down syndrome

	Quality Criteria			*Down Syndrome Study and Sample Characteristics*								*Outcome Data*	
Authors	**Sample**	**Syndrome**	**ASD**	**N**	**% Male**	**Mean Age *(SD) Range***	**% with ID**	**Mean IQ**	**Syndrome Diagnosis**	**ASD Measure**	**ASD Measure Professional**	**ASD Profile**	**%ASD categories**	**% ASD (*N*)**	**Quality Weighting**
Hoffmire *et al.*, 2014				107	53.3	*7-17*	Not reported	Not reported	Clinical[1]	ADI-R	Trained clinical interviewer	No	Autism: 42	42 *(45)*	0.56
Naerland *et al.*, 2017				674	56.7	10.5 *3.1*	Mild:25 Moderate: 56 Severe:21[2]	56.78	Not reported	SCQ	Family-report	Yes	ASD: 37 Autism: 17	17 *(115)*	0.33
Oxelgren *et al.*, 2017				41	70.7	11 *(5-17)*	100	Not reported	Clinical[3]	SCQ ADI-R ADOS DMS-IV/DSM-5	Parent-report Psychologist and special education teacher Team including neuropaediatrician and paediatric nurse	No	SCQ : 26.8 Consensus ADI-R, ADOS, DSM-IV/5: 41	41 *(17)*	0.56
Warner *et al.*, 2014				485	56.29	10.43 *(2.77) 6.0-15.0*	Not reported	Not reported	Confirmed diagnosis[4]	SCQ	Familial report	Yes	ASD: 37.7 Autism: 16.5	16.5 *(80)*	0.44

[1] Participants recruited from a patient registry. Study did not specify whether diagnoses were made by generalists or specialists therefore a conservative rating of one was assigned for 'Syndrome Confirmation'.

[2] Data reported for two cohorts. IQ data based on 106 participants from one cohort (total *n*=175).

[3] The study did not specify whether diagnoses were made by generalists or specialists therefore a conservative rating of one was assigned for 'Syndrome Confirmation'.

[4] The study did not specify whether diagnoses were made by generalists or specialists therefore a conservative rating of one was assigned for 'Syndrome Confirmation'.

4.2.5 Neurofibromatosis

The literature search identified four new studies (all Neurofibromatosis Type 1), although two studies were removed due to insufficient quality ratings (see Table 10). Of the remaining studies, Morris et al. (2016) is notable for its large sample size (*n*=531), however, ratings for Confirmation of Syndrome and ASD Assessment were Adequate only. The study by Plasschaert et al. (2015) generated a high prevalence estimate compared to Richards et al.'s (2015) pooled prevalence estimates in Neurofibromatosis. This study was rated as Excellent for ASD Assessment, basing ASD classification on a detailed multi-disciplinary team assessment using both ADOS-2 (Lord et al., 2012) and DSM-IV-TR criteria. However, it achieved an Adequate rating only for Sample Identification and Confirmation of Syndrome criteria.

Table 10. Quality criteria, participants and study characteristics of research reporting ASD prevalence estimates in Neurofibromatosis

	Quality Criteria		*Neurofibromatosis Study and Sample Characteristics*								*Outcome Data*		
Authors	**Sample Syndrome ASD**	**N**	**% Male**	**Mean Age *(SD)* *Range***	**% with ID**	**Mean IQ**	**Syndrome Diagnosis**	**ASD Measure**	**ASD Measure Professional**	**ASD Profile**	**%ASD categories**	**% ASD (*N*)**	**Quality Weighting**
Constantino *et al.*, 2015		103	Not reported	23 *(17.5)* *3-77*	Not reported	Not reported	Not reported	SRS-2	Self/parent/ spouse/friend-report	Yes	ASD: 12.7	12.7 *(13)*	0.22
Morris *et al.*, 2016		531	46.5	11 *2.5-83.9*	Not reported	Not reported	Clinical[1]	SRS-2	Self/parent/ spouse/friend-report	No	ASD: 13.2	13.2 *(70)*	0.44
Plasschaert *et al.*, 2015		82	54	10.1 *(3.8)* *5-17*	Not reported	Not reported	Clinical[2]	SRS ADOS+ DSM-IV-TR[3]	Parent report MDT	Yes	ASD: 33 ASD: 32.9	32.9 *(27)*	0.56
Plasschaert *et al.*, 2016		42	61.9	12.48 *(3.08)*	Not reported	89.73[4]	Clinical[5]	Not reported	Not reported	No	ASD: 38.1	38.1 *(16)*	0.22

[1] The study did not specify whether diagnoses were made by generalists or specialists therefore a conservative rating of one was assigned for 'Syndrome Confirmation'.
[2] The study did not specify whether diagnoses were made by generalists or specialists therefore a conservative rating of one was assigned for 'Syndrome Confirmation'.
[3] Extensive ASD assessment carried out 'based on clinical suspicion and/or spontaneous complaints by parents and/or teachers'
[4] Based on an abbreviated version of the Dutch Wechsler Intelligence Scale for Children or the Wechsler Adult Intelligence Scale
[5] The study did not specify whether diagnoses were made by generalists or specialists therefore a conservative rating of one was assigned for 'Syndrome Confirmation'

4.2.6 Phenylketonuria

Two new studies were identified, reporting disparate ASD prevalence estimates in individuals with PKU (see Table 11). Bilder et al. (2017) reported a prevalence estimate of 0.7%, drawing on a large dataset (*n*=3714) from healthcare claims databases. This recruitment approach thus relied on ICD-9 codes within the database for PKU and ASD classification. In contrast, Khemir et al. (2016) reported a prevalence estimate of 81%. This study received higher quality ratings for Syndrome Identification (utilising metabolic testing) and ASD Assessment (based on the ADI-R; Lord et al., 1994), however, the study had a much smaller sample of participants recruited from a single paediatric department.

Table 11. Quality criteria, participants and study characteristics of research reporting ASD prevalence estimates in Phenylketonuria

	Quality Criteria			*PKU Study and Sample Characteristics*								*Outcome Data*	
Authors	**Sample**	**Syndrome**	**ASD**	**N**	**% Male**	**Mean Age *(SD) Range***	**% with ID**	**Mean IQ**	**Syndrome Diagnosis**	**ASD Measure**	**ASD Measure Professional**	**ASD Profile**	**%ASD categories**	**% ASD (*N*)**	**Quality Weighting**
Bilder *et al.*, 2017				3714	38.0	38.5 *20-80+*	4.8	Not reported	ICD-9	ICD-9	Not reported	No	ASD: 0.7	0.7 *(27)*	0.44
Khemir *et al.*, 2016				19	46.7	6.5 2-15	Severe: 26.3 Moderate: 26.3 Mild: 10.5[1]	Not reported	Metabolic testing	ADI-R CARS	Not reported	Yes	ADI-R autistic disorder: 81.3 CARS autistic disorder: 78.9	81.3 *(13)*	0.67

[1] Semi-structured evaluation; intellectual disability was assessed based on clinical data in accordance with DSM-5 criteria.

4.2.7 Sotos syndrome

Whilst no studies were identified reporting ASD prevalence estimates in Sotos syndrome in Richards et al.'s (2015) meta-analysis, the current literature search generated two papers (see Table 12). Lane, Milne and Freeth (2017) used a questionnaire method, and thus used family report for syndrome confirmation, however, Sheth et al. (2015) reported a confirmed diagnosis of Sotos syndrome by a clinical geneticist or paediatrician. Both studies relied on a screening tool for assessment of ASD.

Table 12. Quality criteria, participants and study characteristics of research reporting ASD prevalence estimates in Sotos syndrome

	Quality Criteria		*Sotos Syndrome Study and Sample Characteristics*									*Outcome Data*	
Authors	**Sample**	**Syndrome ASD**	**N**	**% Male**	**Mean Age *(SD)* *Range***	**% with ID**	**Mean IQ**	**Syndrome Diagnosis**	**ASD Measure**	**ASD Measure Professional**	**ASD Profile**	**%ASD categories**	**% ASD (*N*)**	**Quality Weighting**
Lane *et al.*, 2017			78	55.1	12.13 *(8.99)* *2.5-50*	Not reported	Not reported	Family-report	SRS-2	Family-report	Yes	ASD: 83.33 *Sub-categories*: Severe clinical range: 55.13 Moderate clinical range: 19.23 Mild clinical range: 8.97	55.13 *(43)*	0.33
Sheth *et al.*, 2015			38	65.8	17.3 *(9.36)* *6.34-43.49*	Not reported	Not reported	Clinical geneticist/ paediatrician	SCQ	Parent/carer report	Yes	ASD: 70.3 Autism: 32.4	32.4 *(12)*	0.56

4.2.8 Angelman syndrome

The literature search identified one new study reporting the prevalence of ASD in individuals with Angelman syndrome (see Table 13). Wink et al. (2015) reported a high prevalence of ASD in their sample (58%) compared to Richards et al.'s (2015) pooled prevalence estimate (34%). This study had a small sample size, but achieved a good overall quality rating. The study confirmed genetic diagnosis of Angelman syndrome in participants and employed both the ADOS (Lord et al., 1989) and ADI-R (Lord et al., 1994) assessment tools. However, they relied on ADOS results only when reporting rates of ASD, therefore the highest ASD Assessment rating could not be assigned.

Table 13. Quality criteria, participants and study characteristics of research reporting ASD prevalence estimates in Angelman syndrome

	Quality Criteria		*Angelman Syndrome Study and Sample Characteristics*									*Outcome Data*	
Authors	**Sample**	**Syndrome ASD**	**N**	**% Male**	**Mean Age *(SD)* *Range***	**% with ID**	**Mean IQ**	**Syndrome Diagnosis**	**ASD Measure**	**ASD Measure Professional**	**ASD Profile**	**%ASD categories**	**% ASD (*N*)**	**Quality Weighting**
Wink *et al.*, (2015)			12	33.3	13.8 *(8.0)* *3-29*	Not reported	18.7 months[1]	Genetic	ADOS ADI-R	Clinicians with expertise in AS	Yes	ADOS ASD: 58.3	58.3 *(7)*	0.78

[1] Mean developmental age based on the cognitive subscale of the Bayley Scales of Infant and Toddler Development (third edition), completed by 11 participants

4.2.9 Noonan syndrome

One new study, by Garg et al. (2017), was identified reporting ASD prevalence estimates in Noonan syndrome (see Table 14). The prevalence estimate for this sample was higher (30%) than the pooled prevalence estimate reported in Richards et al. (2015; 15%). However, this study was of good overall methodological quality. ASD classifications were based on an algorithm developed by the National Institute of Human Development (NICHD) Collaborative Programs of Excellence in Autism (CPEA), which uses results of the ADI-R (Lord et al., 1994), ADOS (Lord et al., 2012) and IQ assessment. Furthermore, Confirmation of Syndrome and Sample Identification quality criteria both received a Good rating.

Table 14. Quality criteria, participants and study characteristics of research reporting ASD prevalence estimates in Noonan syndrome

	Quality Criteria		*Noonan Syndrome Study and Sample Characteristics*										*Outcome Data*	
Authors	**Sample**	**Syndrome ASD**	**N**	**% Male**	**Mean Age *(SD)* *Range***	**% with ID**	**Mean IQ**	**Syndrome Diagnosis**	**ASD Measure**	**ASD Measure Professional**	**ASD Profile**	**%ASD categories**	**% ASD (*N*)**	**Quality Weighting**
Garg *et al.*, 2017			40	62.5	10.83 *(2.75)* *6.0-16.75*	7.7[1]	Not reported	Clinical	ADOS ADI-R	Not reported	Yes	Consensus ASD: 30 Consensus 'broad ASD': 30	30 *(12)*	0.78

[1] Based on the Wechsler Abbreviated Scale of Intelligence-Second edition (39 participants) and the Mullen Scales of Early Learning (one participant).

4.2.10 Cornelia de Lange syndrome

One new study was identified (see Table 15). This study, by Ajmone et al. (2014), received an Excellent rating for Confirmation of Syndrome. However, ASD classification was based on a screening tool only (the Childhood Autism Rating Scale; Schopler, Van Bourgondien, Wellman, & Love, 2010), thus an Adequate rating was assigned. ASD prevalence (35%) was consistent with Richards et al.'s (2015) pooled prevalence estimate, falling within the 95% confidence interval.

Table 15. Quality criteria, participants and study characteristics of research reporting ASD prevalence estimates in Cornelia de Lange syndrome

	Quality Criteria			*Cornelia de Lange Syndrome Study and Sample Characteristics*									*Outcome Data*	
Authors	**Sample**	**Syndrome**	**ASD**	**N**	**% Male**	**Mean Age *(SD)* *Range***	**% with ID**	**Mean IQ**	**Syndrome Diagnosis**	**ASD Measure**	**ASD Measure Professional**	**ASD Profile**	**%ASD categories**	**% ASD (*N*)**	**Quality Weighting**
Ajmone *et al.*, 2014				17	47	8.2 *2.5-13.4*	58.3[1]	Not reported	Molecular analysis	CARS	Clinician	No	Mild autistic features: 24 Severe autistic features: 35	35 *(6)*	0.67

[1] Based on the Leiter International Performance Scale-Revised (eight participants) and the Griffiths Mental Development Scales (four participants).

4.3 Pooled prevalence estimates

Updated pooled prevalence estimates (alongside those reported by Richards et al., 2015, for comparison) are presented for each syndrome in Table 16 (see Appendix 3 for individual forest plots). Confidence intervals (set at 95%) were generated using the Clopper-Pearson method (Clopper & Pearson, 1934). Higgins I^2 values (Higgins, Thompson, Deeks, & Altman, 2003), which provide a measure of the proportion of the observed variance that reflects true difference in ASD prevalence, are presented for each syndrome. I^2 values above 75% indicate high levels of heterogeneity (Higgins et al., 2003), which could be accounted for by study-level covariates. These are highlighted in bold.

Publication bias was assessed by visual inspection of funnel plots, and by calculating Egger's regression test for funnel plot asymmetry (Egger, Smith, & Sterne, 2006)[1]. Where publication bias was indicated, Duval and Tweedie's (2000a; 2000b) 'Trim and Fill' procedure was applied, and a prevalence estimate corrected for publication bias is presented in Table 16.

[1] The power of Egger's regression test is too low to distinguish chance from real asymmetry when study $n<10$.

Table 16. Pooled prevalence estimates, I^2 values and publication bias statistics for each syndrome, ranked according to current quality-weighted pooled prevalence estimates

Syndrome	Number of studies *(number of new studies)*	Number of participants	Random effects model (95% confidence interval)			Quality effects model (95% confidence interval)			Egger's test of asymmetry	Random effects model corrected for publication bias (95% confidence interval)
			2015	Current	I^2 (%)	2015	Current	I^2 (%)		
Rett syndrome	5 *(0)*	194	0.61 (CI: 0.47-0.74)	0.61 (CI: 0.46–0.74)	72.3	0.61 (CI: 0.46-0.74)	0.61 (CI: 0.46–0.74)	72.30	-	-
Cohen syndrome	2 *(0)*	96	0.54 (CI: 0.44-0.64)	0.54 (CI: 0.44–0.64)	0.0	0.54 (CI: 0.44-0.64)	0.54 (CI: 0.44–0.64)	0.0	-	-
Angelman syndrome	8 *(1)*	257	0.35 (CI: 0.24-0.48)	0.37 (CI: 0.26–0.5)	68.5	0.34 (CI: 0.23-0.47)	0.41 (CI: 0.29–0.54)	68.47	-	-
Sotos syndrome	2 *(2)*	116	-	0.44 (CI: 0.23–0.67)	**81.9**	-	0.40 (CI: 0.20–0.65)	**81.89**	-	-
CdLS	13 *(1)*	615	0.43 (CI: 0.33-0.54)	0.41 (CI: 0.32–0.50)	**76.9**	0.43 (CI: 0.32-0.53)	0.40 (CI: 0.31–0.49)	**76.91**	*p*=.443	-
TSC	32 *(7)*	2692	0.37 (CI: 0.33-0.40)	0.34 (CI: 0.29–0.40)	**82.5**	0.36 (CI: 0.33-0.40)	0.34 (CI: 0.29–0.40)	**82.48**	*p*=.069	0.37 (CI: 0.31 – 0.44)
Fragile X	65 *(9)*	5492	0.26 (CI: 0.20-0.31)	0.27 (CI: 0.32–0.32)	**89.7**	0.22 (CI: 0.15-0.30)	0.25 (CI: 0.20–0.29)	**89.66**	*p*=.266	0.32 (CI: 0.27 – 0.38)
CHARGE	4 *(0)*	232	0.29 (CI: 0.14-0.48)	0.25 (CI: 0.10–0.50)	**84.3**	0.30 (CI: 0.14-0.48)	0.22 (CI: 0.08–0.48)	**84.17**	-	0.34 (CI: 0.15 – 0.61)
Noonan syndrome	3 *(1)*	126	0.16 (CI: 0.07-0.27)	0.21 (CI: 0.12–0.33)	52.9	0.15 (CI: 0.07-0.26)	0.20 (CI: 0.11–0.32)	52.87	-	-

Syndrome	Number of studies *(number of new studies)*	Number of participants	Random effects model (95% confidence interval)			Quality effects model (95% confidence interval)			Egger's test of asymmetry	Random effects model corrected for publication bias (95% confidence interval)
			2015	Current	I^2 (%)	2015	Current	I^2 (%)		
NF1	8 *(2)*	1025	0.16 (CI: 0.08-0.26)	0.17 (CI: 0.11–0.26)	**84.8**	0.18 (CI: 0.09-0.29)	0.18 (CI: 0.12–0.27)	**84.79**	-	0.17 (CI: 0.11 – 0.26)
Down syndrome	14 *(4)*	2391	0.16 CI: 0.09-0.23)	0.18 (CI: 0.14–0.24)	**88.5**	0.16 CI: 0.08-0.24)	0.18 (CI: 0.13–0.24)	**88.5**	*p*=.588	0.20 (CI: 0.15 – 0.26)
22q11.2 (VCF)	22 *(8)*	1262	0.12 (CI: 0.06-0.19)	0.27 (CI: 0.24–0.31)	**88.6**	0.11 (CI: 0.05-0.19)	0.15 (CI: 0.09–0.23)	**88.6**	*p*=.0004	0.22 (CI: 0.15 – 0.31)
PKU	5 *(2)*	3997	0.10 (CI: 0.01-0.27)	0.11 (CI: 0.02–0.52)	**97.8**	0.09 (CI: 0.00-0.23)	0.13 (CI: 0.02–0.57)	**97.77**	-	-
Williams syndrome	5 *(0)*	119	0.14 (CI: 0.08-0.21)	0.14 (CI: 0.07–0.24)	32.2	0.12 (CI: 0.06-0.20)	0.11 (CI: 0.05–0.20)	32.16	-	0.16 (CI: 0.09 – 0.25)
Joubert syndrome	2 *(0)*	54	0.09 (CI: 0.00-0.50)	0.08 (CI: 0.00–0.71)	**79.6**	0.09 (CI: 0.00-0.49)	0.09 (CI: 0.00–0.74)	**79.6**	-	-
Moebius syndrome	4 *(0)*	94	0.09 (CI: 0.00-0.25)	0.13 (CI: 0.04–0.33)	57.5	0.07 (CI: 0.00-0.22)	0.06 (CI: 0.01–0.21)	57.51	-	0.23 (CI: 0.07 – 0.52)

Consistent with Richards et al.'s (2015) conclusions, ASD was highly prevalent (prevalence estimates above 30%) in Rett syndrome, Cohen syndrome, tuberous sclerosis complex, Angelman syndrome and Cornelia de Lange syndrome. The current analysis also indicated a high prevalence of ASD in Sotos syndrome (quality-weighted pooled prevalence estimate of 40%), which was not included in Richards et al.'s (2015) analysis due to insufficient literature. However, this estimate was based on only two studies, with a high level of heterogeneity, therefore it should be considered with caution.

ASD was moderately prevalent in individuals with Neurofibromatosis and Fragile X, CHARGE, Noonan, Williams, Down and 22q11.2 deletion syndromes, with quality-weighted pooled prevalence estimates ranging from 11% (Williams syndrome) to 25% (Fragile X syndrome). However, after correcting for publication bias, prevalence estimates for both Fragile X syndrome and CHARGE syndrome were higher than 30%, therefore original estimates may be somewhat conservative.

Pooled prevalence estimates for Moebius syndrome, Joubert syndrome and Phenylketonuria were all deemed not robust by Richards et al. (2015), due to small study numbers and heterogeneity in reported ASD prevalence. The current literature search generated two studies reporting ASD prevalence estimates in individuals diagnosed with Phenylketonuria; although these studies strengthen the pooled prevalence estimate, the meta-analytic result remained highly heterogeneous (I^2=98%) with a broad confidence interval and therefore should be considered with caution. No new studies were identified to strengthen prevalence estimates in Moebius or Joubert syndromes.

Overall, high levels of heterogeneity were indicated, and therefore caution should be applied when interpreting these pooled estimates. I^2 values below 75% were observed for Rett syndrome, Cohen syndrome, Angelman syndrome, Noonan syndrome and Williams syndrome suggesting that, for these syndromes, there was more consistency within the data.

4.3.1 Between-syndrome comparisons

Pooled prevalence estimates for all syndromes, corrected for publication bias where applicable, are presented visually in Figure 3. Relative risk statistics were also calculated from pooled prevalence estimates, to compare the risk of ASD in each syndrome against all other syndromes in the analysis. These are presented in Table 17, with significant relative risk statistics highlighted in bold. Joubert syndrome and Moebius syndrome were not included in these comparisons, as pooled prevalence estimates were considered not robust. In Rett syndrome, Cohen syndrome, Angelman syndrome, Sotos syndrome and Cornelia de Lange syndrome the risk of ASD was significantly higher when compared with at least six other syndromes, indicating that these form a collection of high risk syndromes.

4.3.2 Comparisons between syndromes and the general population

Using the Centre for Disease Control (2014) estimate of ASD prevalence in the general population, odds ratios were generated comparing the odds of ASD in each genetic syndrome with that of the general population. These are presented in Figure 4. Odds ratios ranged from 9.86 (Williams syndrome) to 103.23 (Rett syndrome), evidencing greater likelihood of ASD in all syndromes compared to the general population.

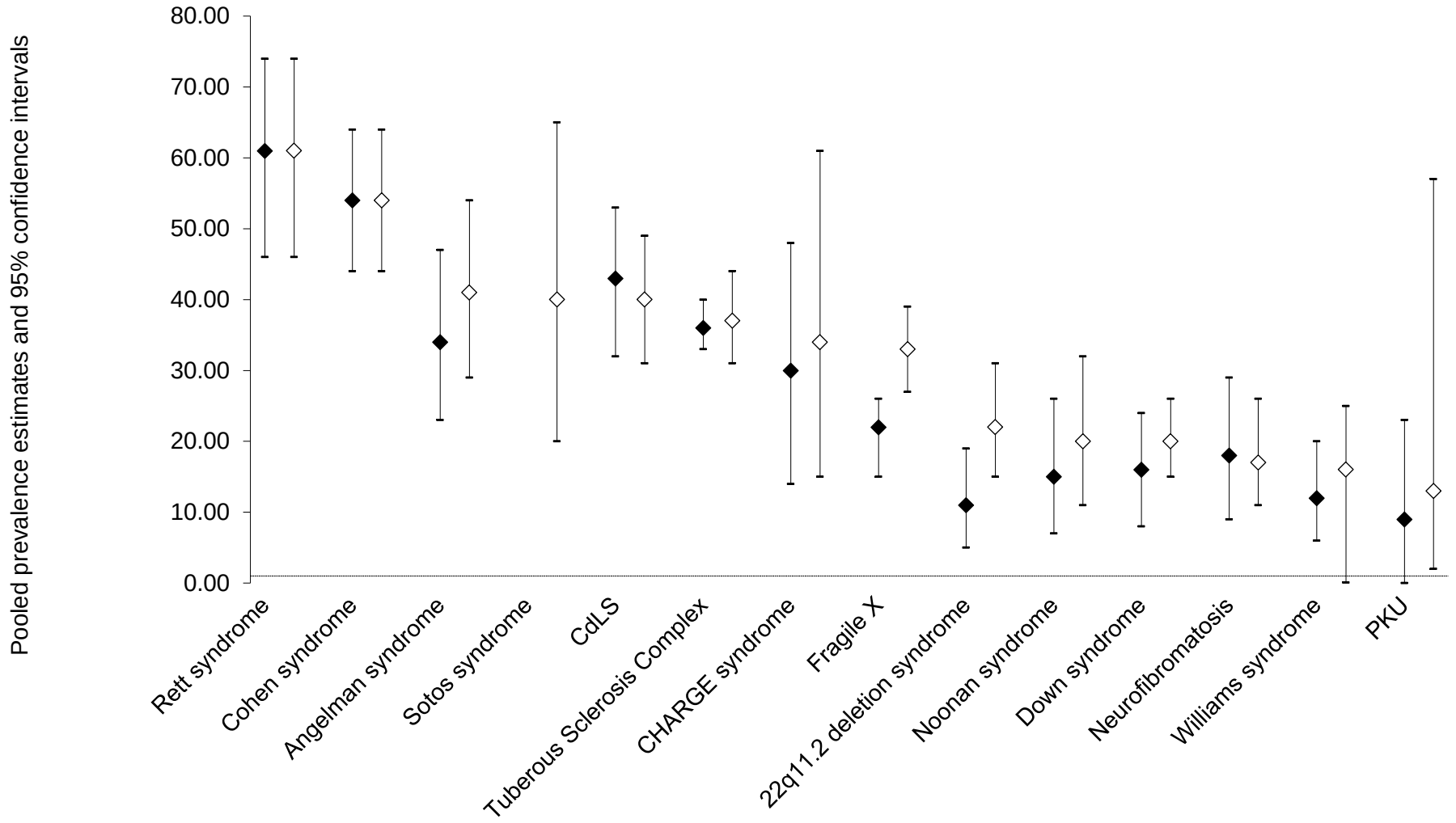

Figure 3. Current pooled prevalence estimates and 95% confidence intervals (unfilled), alongside estimates generated by Richards et al. (2015; filled). Current pooled prevalence estimates corrected for publication bias where applicable.

Table 17. Relative risk statistics with 99% confidence intervals for each syndrome compared against all other syndromes

	Test syndrome													
	Rett	**Cohen**	**AS**	**Sotos**	**CdLS**	**TSC**	**CHARGE**	**FraX**	**22q11.2 (VCF)**	**Noonan**	**DS**	**NF1**	**WS**	**PKU**
Rett	-	0.89 (0.65-1.21)	**0.67 (0.46-0.97)**	**0.66 (0.45-0.96)**	**0.66 (0.45-0.96)**	**0.61 (0.41-0.90)**	**0.56 (0.37-0.84)**	**0.52 (0.34-0.81)**	**0.36 (0.21-0.61)**	**0.33 (0.19-0.57)**	**0.33 (0.19-0.57)**	**0.28 (0.15-0.51)**	**0.26 (0.14-0.49)**	**0.21 (0.11-0.43)**
Cohen	1.13 (0.82-1.55)	-	0.76 (0.51-1.12)	0.74 (0.50-1.10)	0.74 (0.50-1.10)	0.69 (0.45-1.03)	**0.63 (0.41-0.97)**	**0.59 (0.38-0.92)**	**0.41 (0.24-0.70)**	**0.37 (0.21-0.65)**	**0.37 (0.21-0.65)**	**0.31 (0.17-0.58)**	**0.30 (0.16-0.56)**	**0.24 (0.12-0.49)**
AS	1.49 (1.03-2.16)	1.32 (0.89-1.95)	-	0.98 (0.63-1.52)	0.98 (0.63-1.52)	0.90 (0.57-1.43)	0.83 (0.52-1.33)	0.78 (0.48-1.27)	**0.54 (0.30-0.95)**	**0.49 (0.27-0.89)**	**0.49 (0.27-0.89)**	**0.41 (0.22-0.79)**	**0.39 (0.20-0.76)**	**0.32 (0.15-0.66)**
Sotos	1.53 (1.05-2.22)	1.35 (0.91-2.01)	1.03 (0.66-1.60)	-	1.00 (0.64-1.56)	0.93 (0.58-1.47)	0.85 (0.53-1.37)	0.80 (0.49-1.31)	**0.55 (0.31-0.98)**	**0.50 (0.27-0.92)**	**0.50 (0.27-0.92)**	**0.43 (0.22-0.82)**	**0.40 (0.20-0.78)**	**0.33 (0.16-0.68)**
CdLS	1.53 (1.05-2.22)	1.34 (0.91-2.01)	1.03 (0.66-1.60)	1.00 (0.64-1.56)	-	0.93 (0.58-1.47)	0.85 (0.53-1.37)	0.80 (0.49-1.31)	**0.55 (0.31-0.98)**	**0.50 (0.27-0.92)**	**0.50 (0.27-0.92)**	**0.43 (0.22-0.82)**	**0.40 (0.20-0.78)**	**0.33 (0.16-0.68)**
TSC	**1.65 (1.11-2.45)**	1.46 (0.97-2.20)	1.11 (0.70-1.75)	1.08 (0.68-1.72)	1.08 (0.68-1.72)	-	0.92 (0.56-1.50)	0.86 (0.52-1.43)	0.59 (0.33-1.07)	0.54 (0.29-1.00)	0.54 (0.29-1.00)	**0.46 (0.24-0.89)**	**0.43 (0.22-0.85)**	**0.35 (0.17-0.74)**
CHARGE	**1.79 (1.19-2.72)**	**1.59 (1.03-2.44)**	1.21 0.75-1.94	1.18 (0.73-1.90)	1.18 (0.73-1.90)	1.09 (0.67-1.78)	-	0.94 (0.56-1.58)	0.65 (0.35-1.18)	0.59 (0.31-1.10)	0.59 (0.31-1.10)	**0.50 (0.25-0.98)**	**0.47 (0.24-0.94)**	**0.38 (0.18-0.82)**
FraX	**1.91 (1.24-2.93)**	**1.69 (1.08-2.63)**	1.28 0.79-2.09	1.25 (0.76-2.04)	1.25 (0.76-2.04)	1.16 (0.70-1.92)	1.06 (0.63-1.79)	-	0.69 (0.37-1.27)	0.63 (0.33-1.18)	0.63 (0.33-1.18)	0.53 (0.27-1.05)	0.50 (0.25-1.01)	**0.41 (0.19-0.87)**
22q11.2 (VCF)	**2.77 (1.64-4.70)**	**2.45 (1.43-4.22)**	**1.86 (1.05-3.32)**	**1.82 (1.02-3.25)**	**1.82 (1.02-3.25)**	1.68 (0.93-3.04)	1.55 (0.84-2.83)	1.45 (0.79-2.69)	-	0.91 (0.45-1.85)	0.91 (0.45-1.85)	0.77 (0.37-1.63)	0.73 (0.34-1.56)	0.59 (0.26-1.35)
Noonan	**3.05 (1.75-5.32)**	**2.70 (1.53-4.77)**	**2.05 (1.12-3.74)**	**2.00 (1.09-3.66)**	**2.00 (1.09-3.66)**	1.85 (1.00-3.43)	1.70 (0.91-3.19)	1.60 (0.84-3.03)	1.10 (0.54-2.23)	-	1.00 (0.48-2.07)	0.85 (0.39-1.83)	0.80 (0.37-1.75)	0.65 (0.28-1.51)
DS	**3.05 (1.75-5.32)**	**2.70 (1.53-4.77)**	**2.05 (1.12-3.74)**	**2.00 (1.09-3.66)**	**2.00 (1.09-3.66)**	1.85 (1.00-3.43)	1.70 (0.91-3.19)	1.60 (0.84-3.03)	1.10 (0.54-2.23)	1.00 (0.48-2.07)	-	0.85 (0.39-1.83)	0.80 (0.37-1.75)	0.65 (0.28-1.51)
NF1	**3.59 (1.96-6.58)**	**3.18 (1.71-5.89)**	**2.41 (1.26-4.61)**	**2.35 (1.23-4.52)**	**2.35 (1.23-4.52)**	**2.18 (1.12-4.22)**	**2.00 (1.02-3.92)**	1.88 (0.95-3.73)	1.29 (0.61-2.74)	1.18 (0.55-2.54)	1.18 (0.55-2.54)	-	0.94 (0.41-2.14)	0.76 (0.32-1.84)
WS	**3.81 (2.04-7.13)**	**3.38 (1.78-6.38)**	**2.56 (1.31-4.99)**	**2.50 (1.28-4.89)**	**2.50 (1.28-4.89)**	**2.31 (1.17-4.57)**	**2.13 (1.06-4.24)**	2.00 (0.99-4.03)	1.38 (0.64-2.96)	1.25 (0.57-2.74)	1.25 (0.57-2.74)	1.06 (0.47-2.42)	-	0.81 (0.33-1.98)
PKU	**4.69 (2.33-9.44)**	**4.15 (2.04-8.44)**	**3.15 (1.51-6.58)**	**3.08 (1.47-6.44)**	**3.08 (1.47-6.44)**	**2.85 (1.35-6.01)**	**2.62 (1.23-5.58)**	**2.46 (1.14-5.30)**	1.69 (0.74-3.86)	1.54 (0.66-3.58)	1.54 (0.66-3.58)	1.31 (0.55-3.15)	1.23 (0.50-3.00)	-

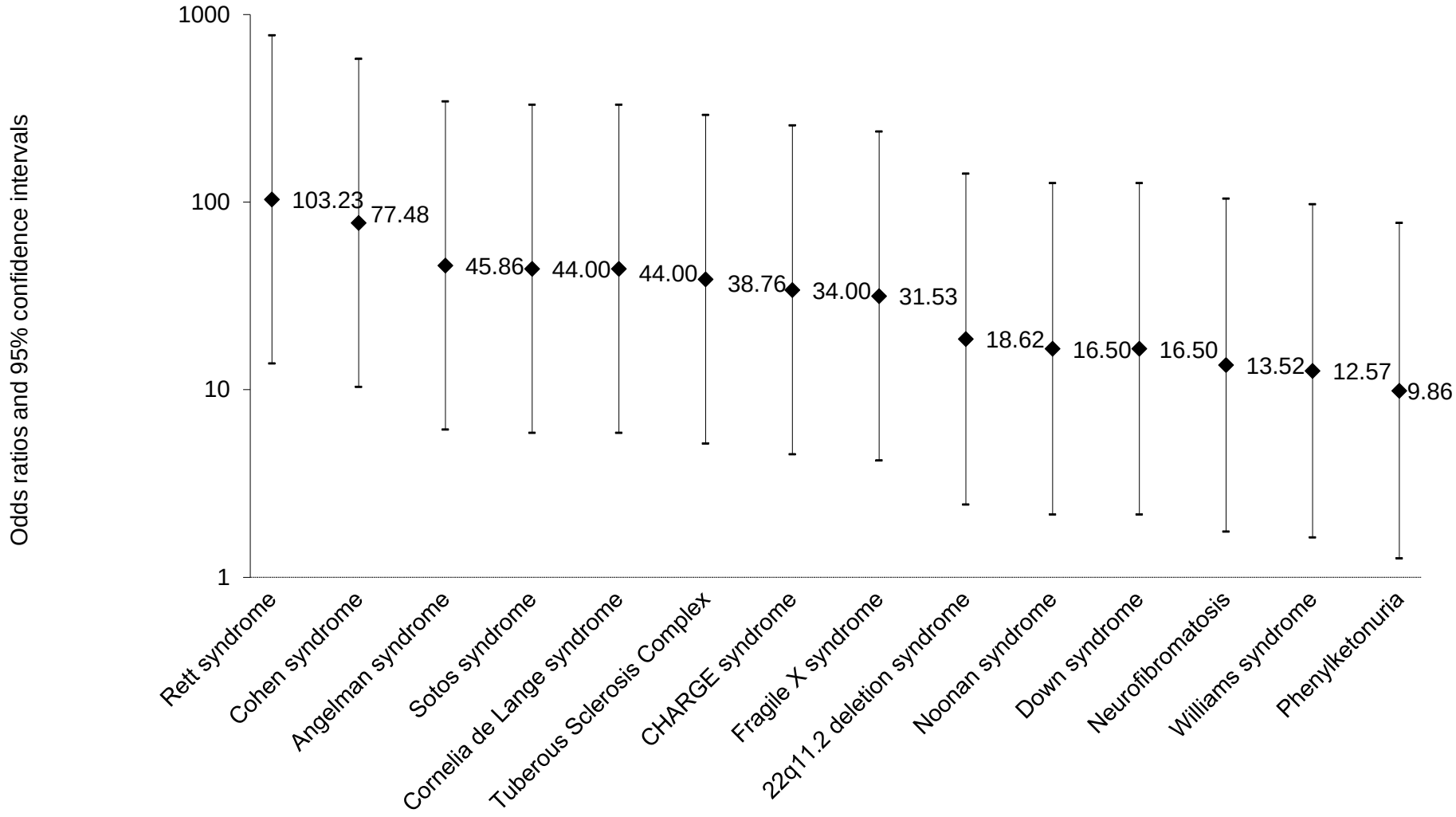

Figure 4. Odds ratios with 95% confidence intervals for each syndrome compared to general population estimates

4.4 Sources of variance

4.4.1 Influential studies

For each syndrome, visual inspection of a Baujat plot (Baujat et al., 2002) and leave-one-out analyses were used to examine the impact of individual studies on the meta-analytic result and data heterogeneity. No studies were identified to be exerting an undue effect on the pooled prevalence estimate for any syndrome group, therefore all studies were retained in the analyses.

4.4.2 Quality criterion and sample characteristics

In order to further explore sources of heterogeneity, sub-group analyses, grouping studies by quality rating on individual quality criterion, were used to evaluate the effect of methodological factors on pooled prevalence estimates. Univariate meta-regression analyses were also conducted to examine the influence of participant characteristics on prevalence estimates. Each analysis was conducted for each syndrome in turn, and subsequently for all syndrome data combined. The results of these analyses are presented in Table 18, with significant effects highlighted in bold. Significant meta-regression analyses are presented with R^2 values, indicating the amount of variance accounted for by the sample characteristic. Where a significant between-group effect of quality criterion rating was observed, individual forest plots displaying pooled prevalence estimates grouped by quality rating are provided in Appendix 4.

The literature searches generated only two studies of sufficient quality reporting ASD prevalence estimates in Cohen syndrome, Sotos syndrome and Joubert syndrome. There

was thus insufficient data to conduct meta-regression analyses in these groups. However, these data were included in subsequent analyses exploring the impact of sample characteristics across all syndrome groups.

Table 18. Results of sub-group and meta-regression analyses evaluating the impact of individual quality criterion ratings and sample characteristics on prevalence estimates. Positive and negative associations are represented by '+' and '-' symbols respectively.

	Study N	RE model	Sub-group analyses: Quality ratings			Meta-regression: Sample characteristics							
						% Male		Mean age (months)		% ID		Mean IQ	
			Sample Identification	Confirmation of Syndrome	ASD Assessment	*Study N*	*p value*	*Study N*	*p value*	*Study N*	*p value*	*Study N*	*p value*
Angelman syndrome	8	0.37 (CI: 0.26–0.5) I^2=68.5	*p*=.725	***p*=.026**	*p*=.084	8	*p*=.116	5	*p*=.800	1	-	0	-
CdLS	13	0.41 (CI: 0.32–0.50) I^2=**76.9**	*p*=.58	*p*=.20	***p*=.0007**	13	*p*=.501	12	*p*=.405	9	*p*=.596	0	-
CHARGE syndrome	4	0.25 (CI: 0.10-0.50) I^2=**84.3**	*p*=.45	*p*=.80	*p*=.97	3	***p*=.0267 (-) R^2=79.67**	3	*p*=.773	2	-	0	-
Cohen syndrome	2	0.54 (CI: 0.44–0.64) I^2=0.0	Both studies rated Good	Both studies rated Adequate	Both studies rated Excellent	2	-	2	-	2	-	2	-
Down syndrome	14	0.18 (CI: 0.14–0.24) I^2=**88.5**	***p*<.0001**	*p*=.818	**p<.0001**	12	*p*=.913	8	*p*=.903	3	All 100%	1	-
Fragile X syndrome	65	0.27 (CI: 0.23-0.32) I^2=**89.7**	*p*=.133	***p*<.0001**	*p*=.566	65	*p*=.232	46	***p*=.025 (-) R^2=0.0**	18	*p*=.655	25	*p*=.123
Joubert syndrome	2	0.08 (CI: 0.00–0.71) I^2=**79.6%**	Both studies rated Good	***p*=.027**	***p*=.027**	2	-	2	-	0	-	0	-
Moebius syndrome	4	0.13 (CI: 0.04–0.33) I^2=57.5	Both studies rated Good	*p*=.411	***p*=.012**	2	-	2	-	2	-	1	-
NF1	8	0.17 (CI: 0.11–0.26) I^2=**84.8**	***p*<.0001**	*p*=.586	***p*=.001**	7	*p*=.826	6	*p*=.965	1	-	2	-

	Study N	RE model	Sub-group analyses: Quality ratings			Meta-regression: Sample characteristics							
			Sample Identification	Confirmation of Syndrome	ASD Assessment	% Male		Mean age (months)		% ID		Mean IQ	
						Study N	*p value*	*Study N*	*p value*	*Study N*	*p value*	*Study N*	*p value*
Noonan syndrome	3	0.21 (CI: 0.12–0.33) I^2=52.9	All studies rated Good	*p*=.07	*p*=.127	3	*p*=.61	1	-	1	-	1	-
PKU	5	0.11 (CI: 0.02–0.52) I^2=**97.8**	***p*<.0001**	***p*<.0001**	***p*=.0001**	3	*p*=.364	4	***p*=.001 (-) R^2=81.73**	5	*p*=.15	1	-
Rett syndrome	5	0.61 (CI: 0.46–0.74) I^2=72.3	***p*=.0096**	*p*=.233	***p*=.001**	5	All 0%	2	-	1	-	0	-
Sotos syndrome	2	0.44 (CI: 0.23–0.67) I^2=**81.9**	Both studies rated Good	***p*=.019**	Both studies rated Adequate	2	-	2	-	0	-	0	-
TSC	32	0.34 (CI: 0.29–0.40) I^2=**82.5**	*p*=.106	*p*=.56	***p*=.0026**	24	*p*=.279	17	*p=.450*	16	*p*=.112	4	*p*=.602
22q11.2 deletion syndrome	22	0.27 (CI: 0.24–0.31) I^2=**88.6**	*p*=.142	All studies rated Excellent	***p*<.0001**	21	*p*=.182	13	*P=.269*	5	*p*=.694	13	***p*=.0132 (-) R^2=36.75**
Williams syndrome	5	0.14 (CI: 0.07–0.24) I^2=32.2	*p*=.357	***p*=.017**	*p*=.808	5	*p*=.245	5	***p*=.020 (+) R^2=100**	1	-	3	*p*=.198
All data	194	-	*p*=.441	*p*=.214	*p*=.194	177	*p*=.221	130	***p*=.014 (-) R^2=0.53**	67	***p*=.001 (+) R^2=22.01**	53	***p*=.04 (-) R^2=1.09**

Analysis of individual quality criteria revealed that ASD prevalence estimates were significantly different according to Sample Identification ratings in four syndromes (Down syndrome, Neurofibromatosis, Phenylketonuria and Rett syndrome), $p<.0001$ – $p=.0096$. ASD prevalence estimates were significantly different dependent on Confirmation of Syndrome ratings in six syndromes (Angelman syndrome, Fragile X syndrome, Joubert syndrome, Phenylketonuria, Sotos syndrome and Williams syndrome), $p<.0001$ – $p=.027$. Finally, prevalence estimates were significantly different according to ASD Assessment ratings in nine syndromes (Cornelia de Lange syndrome, Down syndrome, Joubert syndrome, Moebius syndrome, Neurofibromatosis, Phenylketonuria, Rett syndrome, tuberous sclerosis complex and 22q11.2 deletion syndrome), $p<.0001$ – $p=.027$. Thus, in individual syndromes methodological factors exerted a significant influence on ASD prevalence estimates. However, when all syndrome data were combined, no individual quality criterion had a significant effect on ASD prevalence estimates (all $p>.05$).

The proportion of males in the sample significantly predicted ASD prevalence estimates in CHARGE syndrome only ($p=.027$). In CHARGE syndrome a negative association was observed, accounting for 79.67% of the variance. However, this analysis was based on three papers only, and therefore should be interpreted with caution. When all syndrome data were combined, the association between sample gender ratios and prevalence estimates was non-significant ($p>.05$).

The mean age of participants was significantly positively associated with ASD prevalence estimates in Williams syndrome ($p=.02$), but was negatively associated with ASD prevalence estimates in Fragile X syndrome ($p=.025$) and Phenylketonuria ($p=.001$). Mean age was also negatively associated with ASD prevalence when all syndrome data were

combined (p=.014), suggesting that, across all syndromes, ASD phenomenology decreased with age. However, it is important to note that although this was statistically significant, mean age accounted for a small proportion of the variance in ASD prevalence estimates (0.53%).

The influence of intellectual ability was assessed by evaluating the predictive value of both mean IQ and the proportion of the sample with intellectual disability. Only four syndromes had sufficient data to analyse the influence of mean IQ, with a significant effect identified in only one syndrome; mean IQ negatively predicted ASD prevalence estimates in 22q11.2 deletion syndrome (p=.013). The influence of the proportion of the sample with intellectual disability on ASD prevalence was evaluated for five syndromes, however, no significant associations were identified. When all syndrome data were combined, intellectual disability was significantly positively associated with ASD prevalence (p=.001), accounting for 22.01% of the variance. Mean IQ was negatively associated with ASD prevalence (p=.04), but accounted for a smaller proportion of the variance (1.09%).

5. Discussion

This study aimed to update Richards et al.'s (2015) meta-analysis, synthesising the most recent data to provide robust estimates of ASD prevalence in rare genetic syndromes. The current estimates were strengthened by calculations using standardised and auditable algorithms, which have been judged valid through peer review. Given significant heterogeneity in ASD prevalence estimates, the current study reported further analyses exploring the effect of methodological factors and sample characteristics on ASD prevalence estimates. Results revealed that methodological issues, such as sample recruitment, confirmation of syndrome diagnoses, and ASD assessment methods all have the potential to significantly affect ASD prevalence estimates. Furthermore, when all syndrome data were combined prevalence estimates increased as the mean age and IQ of the sample decreased and the proportion of participants with intellectual disability increased. These analyses add valuable new insights into sources of variance in ASD prevalence estimates in syndrome groups, and therefore the extent to which heightened ASD phenomenology in these syndromes reflects 'true' syndrome-specific factors.

5.1 Updated prevalence estimates

The current literature search identified 37 new studies that reported ASD prevalence estimates in rare genetic syndromes and met pre-defined quality criteria. Overall, prevalence estimates remained consistent with those reported in Richards et al. (2015), with high prevalence estimates generated for Rett syndrome, Cohen syndrome, tuberous sclerosis complex, Angelman syndrome and CHARGE syndrome (all >30%). Prevalence estimates in Sotos syndrome were not reported in Richards et al.'s (2015) analysis, due to

an insufficient number of studies; the current search identified two studies that produced a high pooled prevalence estimate of 40%. ASD was moderately prevalent in Neurofibromatosis, Phenylketonuria and Fragile X, CHARGE, Noonan, Williams, Down and 22q11.2 deletion syndromes, with quality-weighted pooled prevalence estimates ranging from 11% (Williams syndrome) to 25% (Fragile X syndrome). However, after correcting for publication bias, prevalence estimates for both Fragile X syndrome and CHARGE syndrome were higher than 30%. Original prevalence estimates for Fragile X syndrome and CHARGE syndrome may, therefore, be somewhat conservative and the present update provides an essential revision. Importantly, odds ratios showed that in *all* syndromes, the odds of ASD were higher than that of the general population, based on the Center for Disease Control (2014) estimate.

Lee et al. (2016) suggest that genetic syndromes evidencing heightened ASD phenomenology offer a "simplified context" within which to explore links between genetics, neurobiology, cognition and behaviour in ASD. This approach could offer new models for understanding the complex biological pathways underpinning nonsyndromic ASD. The current findings provide robust prevalence estimates, indicating those syndromes in which the prevalence of ASD is most consistently reported to be high, thus helping to focus such research efforts.

5.2 Sources of variance: Methodological factors

Significant heterogeneity in prevalence estimates were reported both within and across syndromes. To explore the sources of this variance, additional analyses were conducted to evaluate the influence of methodological factors on ASD prevalence estimates. These

analyses revealed that ASD prevalence estimates were significantly different according to Confirmation of Syndrome ratings in six syndromes (Angelman syndrome, Fragile X syndrome, Joubert syndrome, Phenylketonuria, Sotos syndrome and Williams syndrome), and according to Sample Identification ratings in four syndromes (Down syndrome, Neurofibromatosis, Phenylketonuria and Rett syndrome). These criteria rated studies based on the methods employed to establish syndrome diagnoses (i.e. genetic/molecular, clinical, or unspecified), and the sample recruitment methods (i.e. random or total population samples, participants recruited from multiple or single restricted sites, or unspecified sampling strategies). In many syndromes more than one genetic variant exists, and clinical diagnoses can be determined by multiple genetic and behavioural criteria. It is possible, therefore, that ASD phenomenology will vary according to within-syndrome genetic and behavioural variation (Howlin, Karpf, & Turk, 2005), and therefore according to the method through which studies confirm syndrome diagnoses. In contrast, one might hypothesise that differences in ASD prevalence estimates according to Sample Identification ratings reflect biases inherent in sampling strategies. Recruitment from clinics or syndrome-linked associations, for example, might plausibly be biased toward individuals with additional behavioural and psychological difficulties, potentially generating higher ASD prevalence estimates.

Due to low study numbers, many of the within-syndrome analyses were based on a small number of comparisons. As a result, patterns of association between prevalence estimates and methods of recruitment and syndrome confirmation were not clear. Nonetheless, these preliminary findings indicate that these methodological factors have the potential to significantly influence ASD prevalence estimates, highlighting the importance of robust methodologies. The quality of sample identification methods in particular remains an issue,

with only nine (4.6%) studies achieving an Excellent rating. This is arguably reasonable given the low prevalence of these syndromes in the general population, and therefore a reliance on opportunity samples, however, future studies should focus on ensuring that research samples are representative of the entire syndrome group.

A significant effect of ASD assessment method was observed in nine syndromes (Cornelia de Lange syndrome, Down syndrome, Joubert syndrome, Moebius syndrome, Neurofibromatosis, Phenylketonuria, Rett syndrome, tuberous sclerosis complex, and 22q11.2 deletion syndrome). In these groups, studies classifying ASD estimates based on a consensus across multiple measures, including gold-standard diagnostic tools, consistently produced more conservative estimates than those relying on a single measure of ASD. More broadly, Moss and Howlin (2009) suggest that it is more challenging to detect behaviours *specific* to ASD in the context of intellectual disability; given that levels of intellectual disability are high across all syndrome groups assessed, detailed measures offering greater specificity are likely to provide more reliable estimates when compared to screening instruments or parental report. However, it is important to note that even gold-standard diagnostic instruments, such as the ADOS (Lord et al., 2012) and ADI-R (Lord et al., 1994), might have limited sensitivity and specificity in individuals with intellectual disability (Gray et al., 2008; Sappok et al., 2013). Abbeduto, McDuffie and Thurman (2014) suggest that these diagnostic tools similarly offer insufficient sensitivity to distinguish atypical development observed in Fragile X syndrome, for example, from atypical development *specific* to ASD. Together with these studies the current findings, which evidence the influence of ASD assessment method on prevalence estimates, highlight the need for further research developing the validity of ASD assessment tools in those with intellectual disability and identified genetic syndromes. Findings by Gotham et

al. (2007) illustrate the potential for such developments, demonstrating improved sensitivity and specificity of the ADOS when different algorithms were generated for sub-samples grouped by age and verbal ability. Such an approach could similarly be extended to determine optimal behavioural coding schemes for individuals with particular syndrome diagnoses and intellectual disability.

5.3 Sources of variance: Sample characteristics

An exploration of sample characteristics revealed that the proportion of male participants significantly predicted variance in ASD prevalence estimates in CHARGE syndrome only. In this syndrome a higher proportion of male participants was associated with lower ASD prevalence estimates. However, this analysis was based on three papers only, and should therefore be interpreted with caution. When all syndrome data were combined, the gender ratio did not significantly predict prevalence estimates. Given that a male to female ratio of 4:1 is observed in nonsyndromic ASD (Baird et al., 2006), it was hypothesised that higher ASD prevalence estimates in syndromes with significantly skewed gender ratios (such as Fragile X syndrome) could, to some extent, be explained by a higher proportion of affected males. The current findings, however, suggest that the influence of gender on heightened ASD prevalence estimates in these syndromes is likely to be minimal.

Consistent with evidence in nonsyndromic ASD (Seltzer et al., 2004; Woodman et al., 2015), when all syndrome data were combined ASD prevalence significantly decreased as the mean age of participants increased, although age accounted for little variance in prevalence estimates. Mean age was positively associated with ASD prevalence estimates in Williams syndrome, but negatively associated with ASD prevalence estimates in

Phenylketonuria and Fragile X syndrome. This is discrepant with previous studies demonstrating positive associations between age and ASD severity in children and adolescents with Fragile X syndrome (Hatton et al., 2006; Lee et al., 2016). One possible explanation for these contradictory findings is that the relationship between age and ASD phenomenology is non-linear. In this case, different trends would be observed over early childhood, adolescence and adulthood. The current analyses included widely varying age groups and studies with broad age ranges, potentially obscuring more nuanced relationships between age and ASD prevalence. Additionally, age-related changes might be different within each domain of ASD phenomenology. In adults and adolescents with non-syndromic ASD, for example, Woodman et al. (2015) reported particular improvements in verbal communication and restricted and repetitive behaviours. It is possible that the current analyses, which focused on an overall ASD classification, were not sensitive to changes in the *profile* of ASD symptomatology.

In nine syndromes IQ and intellectual disability data were each reported by fewer than three papers, therefore meta-regression analyses were not possible. However, mean IQ negatively predicted ASD estimates in 22q11.2 deletion syndrome. When all syndrome data were combined, the proportion of the sample with intellectual disability significantly positively predicted ASD prevalence, accounting for 22% of the variance. Furthermore, mean IQ was negatively associated with ASD prevalence, although this accounted for a smaller proportion of the variance (1.09%). These findings add to existing evidence for a negative association between intellectual ability and ASD phenomenology in Fragile X syndrome, Angelman syndrome and Down syndrome (Hall et al., 2010; Loesch et al., 2007; Molloy et al., 2009; Thurman et al., 2015; Trillingsgaard & Østergaard, 2004). Skuse (2007) suggests that intellectual disability might limit an individual's capacity to

compensate for autistic vulnerabilities, and therefore acts as an additional risk factor in the development of ASD-linked behaviour. As discussed above, lower sensitivity and specificity of ASD assessment tools for individuals with intellectual disability might also conflate ASD prevalence estimates in genetic syndromes associated with more severe levels of intellectual disability (Gray et al., 2008; Moss & Howlin, 2009; Sappok et al., 2013). Together these findings are consistent with the notion that the link between genetic syndromes and ASD phenomenology is, to some extent, accounted for by the degree of intellectual disability associated with these syndromes (Moss & Howlin, 2009).

5.4 Clinical implications

The findings discussed above have several clinical implications. Pooled prevalence estimates indicated that ASD phenomenology was higher in *all* syndromes assessed when compared to general population estimates. However, in clinical settings diagnostic overshadowing, whereby professionals are less sensitive to a range of potentially distinct behavioural or psychological processes in the context of another, existing diagnostic label (Mason & Scior, 2004), might prevent the recognition of additional, ASD-linked difficulties in individuals with a diagnosed genetic syndrome (Moss & Howlin, 2009). In such cases, individuals are unlikely to be provided with appropriately targeted clinical and educational intervention. Given the emphasis on early diagnosis and intervention for improving outcomes in individuals with ASD (Kim & Lord, 2012), the current findings highlight the need for professionals to recognise behaviours specific to ASD presentations in these syndrome groups, and the importance of providing ASD-specific assessment and intervention as required.

The current findings also suggest caution is required when interpreting ASD diagnostic tools in these groups. Analyses revealed a significant effect of ASD assessment method on prevalence estimates in nine syndromes, which, along with previous studies, highlighted some of the issues with ASD assessment measures in individuals with genetic syndromes and intellectual disability. In the absence of diagnostic tools specifically designed for these populations, ASD diagnoses should be based on a comprehensive assessment and careful formulation, conducted by professionals with specialist knowledge of the cognitive and behavioural profiles typical of these genetic syndromes.

5.5 Limitations and future research

The current study provides the most up to date, robust ASD prevalence estimates in genetic syndromes associated with elevated rates of ASD. These estimates are strengthened by the use of standardised and auditable algorithms for calculating pooled prevalence estimates. Given significant heterogeneity in prevalence estimates, the current study added valuable analyses of predictor variables, evidencing the influence of methodological factors and sample characteristics on ASD prevalence estimates.

Despite these strengths, a number of limitations are worthy of consideration. First, the number of studies available for some syndrome groups remained small, and many studies failed to report participant characteristics. As a result, many within-syndrome meta-regression analyses were not possible or were based on a small number of studies. These results therefore require replication with larger study numbers as the available literature expands. IQ data in particular was limited; only 67 (35%) studies reported the proportion of the sample with intellectual disability, and only 53 (27%) reported participants' mean

IQ. Given the findings presented above, which indicate the influence of these variables on ASD prevalence estimates, a measure of intellectual ability should be clearly reported in future research.

Interpretations were also limited to univariate meta-regression analyses. Sample characteristics such as age and IQ likely co-vary, and may be differentially associated with different syndromes. In order to account for these interactions, multivariate analyses would require complex models and therefore very large datasets. Insufficient study numbers and missing data prevented these more nuanced, multivariate meta-regression analyses, thus it was not possible to assess the *relative* influence of sample characteristics against syndrome classification[1]. With increasing study numbers, future research should delineate the interactions between age, IQ, syndrome diagnoses and ASD phenomenology.

The current study maintained Richards et al.'s (2015) quality criteria, weighting studies according to key methodological factors. These broad criteria were developed with the purpose of excluding weak studies whilst retaining sufficient data for a comprehensive meta-analysis. However, given that methodological factors had a significant effect on ASD prevalence estimates in individual syndrome analyses, future reviews would benefit from applying more rigorous criteria in order to examine prevalence estimates generated by the highest quality research. Despite these limitations, the decision to apply Richards et al.'s (2015) criteria in the current study allowed for comparisons with earlier prevalence estimates and enabled the inclusion of sufficient data for further analyses examining sources of variance.

[1] A simple multivariate meta-regression model was calculated and results are presented in Appendix 5. However, this simple model was judged not meaningful and was therefore excluded from the report.

Finally, the current analyses did not consider ASD symptom profiles, instead focusing on categorical classifications. However, evidence increasingly suggests that individuals with a range of genetic syndromes, including Fragile X syndrome, Rett syndrome, Cohen syndrome and Cornelia de Lange syndrome, display distinct profiles of strength and impairment within ASD diagnostic criteria (Abbeduto et al., 2014; Howlin et al., 2005; Moss, Oliver, Nelson, Richards, & Hall, 2013). Cornish et al. (2007), for example, suggest that individuals with Fragile X syndrome are socially motivated, with a good understanding of social interaction, but display gaze avoidant behaviour due to social anxiety and hypersensitivity. Individuals with Rett syndrome, in contrast, show preserved eye contact despite impairments in other areas of social interaction (Nomura & Segawa, 2005). Differing symptom profiles might indicate different underlying neurological or psychological mechanisms. This has implications for understanding the gene-cognition-behaviour links in genetic syndromes and ASD, and for providing appropriately targeted clinical intervention (Moss & Howlin, 2009). The current findings therefore provide estimates of the prevalence of ASD as defined by existing diagnostic guidance and assessment tools. However, the extent to which this construct equates to ASD phenomenology observed in individual syndromes requires further research.

5.6 Conclusion

In conclusion, the current study provided updated, robust estimates of ASD prevalence in rare genetic syndromes. These estimates help to guide future research into gene-behaviour links in ASD and highlight the need for appropriately targeted assessment and intervention in individuals with syndrome diagnoses. Analyses demonstrating the impact of methodological factors on ASD prevalence estimates emphasised the importance of robust

methodologies and appropriate tools for assessing ASD in individuals with intellectual disability and genetic syndromes. Meta-regression analyses showed an association between ASD prevalence and intellectual disability, adding weight to the argument that intellectual disability plays a role in heightened ASD phenomenology in genetic syndromes. As research in genetic syndromes develops, future studies should aim to further delineate the relative influence of sample variables and syndrome-specific factors on ASD symptom profiles.

6. References

Abbeduto, L., McDuffie, A., & Thurman, A. J. (2014). The Fragile X syndrome–autism comorbidity: What do we really know? *Frontiers in Genetics, 5,* 1-10.

Adlof, S. M., Klusek, J., Shinkareva, S. V., Robinson, M. L., & Roberts, J. E. (2015). Phonological awareness and reading in boys with Fragile X syndrome. *Journal of Child Psychology and Psychiatry, and Allied Disciplines, 56*(1), 30–39.

Ajmone, P. F., Rigamonti, C., Dall'Ara, F., Monti, F., Vizziello, P., Milani, D., …& Costantino, A. (2014). Communication, cognitive development and behavior in children with Cornelia de Lange Syndrome (CdLS): Preliminary results. *American Journal of Medical Genetics. Part B, Neuropsychiatric Genetics, 165B*(3), 223–229.

American Psychiatric Association (2000). *Diagnostic and Statistical Manual of Mental Disorders*. Washington, DC: Author.

American Psychiatric Association (2013). *Diagnostic and Statistical Manual of Mental Disorders.* (5th ed.). Arlington, VA: American Psychiatric Publishing.

Baird, G., Simonoff, E., Pickles, A., Chandler, S., Loucas, T., Meldrum, D., & Charman, T. (2006). Prevalence of disorders of the autism spectrum in a population cohort of children in South Thames: The Special Needs and Autism Project (SNAP). *The Lancet, 368*(9531), 210–215.

Barendregt, J. J., & Doi, S. A. (2011). Meta XL User Guide. Version 2.0.

Baron-Cohen, S., Scott, F. J., Allison, C., Williams, J., Bolton, P., Matthews, F. E., & Brayne, C. (2009). Prevalence of autism-spectrum conditions: UK school-based population study. *The British Journal of Psychiatry: The Journal of Mental Science*, *194*(6), 500–509.

Baujat, B., Mahé, C., Pignon, J., & Hill, C. (2002). A graphical method for exploring heterogeneity in meta-analyses: Application to a meta-analysis of 65 trials. *Statistics in Medicine*, *21*(18), 2641–2652.

Betancur, C., & Buxbaum, J. D. (2013). SHANK3 haploinsufficiency: A 'common' but underdiagnosed highly penetrant monogenic cause of autism spectrum disorders. *Molecular Autism*, *4*(1), 17.

Bilder, D. A., Kobori, J. A., Cohen-Pfeffer, J. L., Johnson, E. M., Jurecki, E. R., & Grant, M. L. (2017). Neuropsychiatric comorbidities in adults with Phenylketonuria: A retrospective cohort study. *Molecular Genetics and Metabolism, 121*(1), 1–8.

Center for Disease Control (2014). Prevalence of autism spectrum disorders - Autism and Developmental Disabilities Monitoring Network, United States, 2010. *MMWR Suveillance*, *63*, 1–22.

Clopper, C., & Pearson, S. (1934). The use of confidence or fiducial limits illustrated in the case of the binomial. *Biometrika*, *26*, 404–413.

Constantino, J. N., Zhang, Y., Holzhauer, K., Sant, S., Long, K., Vallorani, A., …& Gutmann, D. H. (2015). Distribution and within-family specificity of quantitative autistic traits in patients with Neurofibromatosis type I. *The Journal of Pediatrics*, *167*(3), 621–626.

Cornish, K., Turk, J., & Levitas, A. (2007). Fragile X syndrome and autism: Common developmental pathways? *Current Paediatric Reviews, 3*(1), 61-68.

Crawford, D. C., Acuña, J. M., & Sherman, S. L. (2001). FMR1 and the Fragile X syndrome: Human genome epidemiology review. *Genetics in Medicine, 3*(5), 359–371.

de Sonneville, L. M. J., Hidding, E., van Engeland, H., Vorstman, J. A. S., Sijmens-

Morcus, M. E. J., & Swaab, H. (2018). Executive functioning and its relation to ASD and ADHD symptomatology in 22q11.2 deletion syndrome. *Child Neuropsychology*, *24*(1), 1–19.

de Vries, P. J., Hunt, A., & Bolton, P. F. (2007). The psychopathologies of children and adolescents with tuberous sclerosis complex (TSC): A postal survey of UK families. *European Child & Adolescent Psychiatry*, *16*(1), 16–24.

Duval, S., & Tweedie, R. (2000a). A nonparametric 'trim and fill' method of accounting for publication bias in meta-analysis. *Journal of the American Statistical Association*, *95*(449), 89–98.

Duval, S., & Tweedie, R. (2000b). Trim and fill: A simple funnel-plot-based method of testing and adjusting for publication bias in meta-analysis. *Biometrics*, *56*(2), 455–463.

Egger, M., Smith, G., & Sterne, J. (2006). Bias in meta-analysis detected by a simple, graphical test. *Statistics in Medicine*, *25*, 3443–4357.

Fiksinski, A. M., Breetvelt, E. J., Duijff, S. N., Bassett, A. S., Kahn, R. S., & Vorstman, J. A. S. (2017). Autism spectrum and psychosis risk in the 22q11.2 deletion syndrome: Findings from a prospective longitudinal study. *Schizophrenia Research*, *188*, 59–62.

Fjermestad, K. W., Vatne, T. M., & Gjone, H. (2015). Cognitive behavioral therapy for adolescents with 22q11.2 deletion syndrome. *Advances in Mental Health and Intellectual Disabilities*, *9*(1), 30–39.

Garg, S., Brooks, A., Burns, A., Burkitt-Wright, E., Kerr, B., Huson, S., …& Green, J. (2017). Autism spectrum disorder and other neurobehavioural comorbidities in rare

disorders of the Ras/MAPK pathway. *Developmental Medicine and Child Neurology*, *59*(5), 544–549.

Garg, S., Plasschaert, E., Descheemaeker, M-J., Huson, S., Borghgraef, M., Vogels, A., …& Green, J. (2015). Autism spectrum disorder profile in Neurofibromatosis type I. *Journal of Autism and Developmental Disorders*, *45*(6), 1649–1657.

Gotham, K., Risi, S., Pickles, A., & Lord, C. (2007). The Autism Diagnostic Observation Schedule: Revised algorithms for improved diagnostic validity. *Journal of Autism and Developmental Disorders*, *37*(4), 613–627.

Gray, K. M., Tonge, B. J., & Sweeney, D. J. (2008). Using the Autism Diagnostic Interview-Revised and the Autism Diagnostic Observation Schedule with young children with developmental delay: Evaluating diagnostic validity. *Journal of Autism and Developmental Disorders*, *38*(4), 657–667.

Grefer, M., Flory, K., Cornish, K., Hatton, D., & Roberts, J. (2016). The emergence and stability of attention deficit hyperactivity disorder in boys with Fragile X syndrome. *Journal of Intellectual Disability Research*, *60*(2), 167–178.

Greiss Hess, L., Fitzpatrick, S. E., Nguyen, D. V., Chen, Y., Gaul, K. N., Schneider, A., …& Hagerman, R. J. (2016). A randomized, double-blind, placebo-controlled trial of low-dose sertraline in young children with Fragile X syndrome. *Journal of Developmental and Behavioral Pediatrics*, *37*(8), 619–628.

Hall, S. S., Lightbody, A. A., Hirt, M., Rezvani, A., & Reiss, A. L. (2010). Autism in Fragile X syndrome: A category mistake? *Journal of the American Academy of Child and Adolescent Psychiatry*, *49*(9), 921–933.

Hall, S. S., & Venema, K. M. (2017). A screening tool to measure eye contact avoidance in boys with Fragile X syndrome. *Journal of Autism and Developmental Disorders*,

47(7), 2254–2264.

Hartley, S. L., Wheeler, A. C., Mailick, M. R., Raspa, M., Mihaila, I., Bishop, E., & Bailey, D. B. (2015). Autism symptoms across adulthood in men with Fragile X syndrome: A cross-sectional analysis. *Journal of Autism and Developmental Disorders*, *45*(11), 3668–3679.

Hartshorne, N., Hudson, A., MacCuspie, J., Kennert, B., Nacarato, T., Hartshorne, T., & Blake, K. (2016). Quality of life in adolescents and adults with CHARGE syndrome. *American Journal of Medical Genetics. Part A*, *170*(8), 2012–2021.

Hatton, D. D., Sideris, J., Skinner, M., Mankowski, J., Bailey, D. B., Roberts, J., & Mirrett, P. (2006). Autistic behavior in children with Fragile X syndrome: Prevalence, stability, and the impact of FMRP. *American Journal of Medical Genetics Part A*, *140A*(17), 1804–1813.

Hidding, E., Swaab, H., de Sonneville, L. M. J., van Engeland, H., Sijmens-Morcus, M. E. J., Klaassen, P. W. J., …& Vorstman, J. A. S. (2015). Intellectual functioning in relation to autism and ADHD symptomatology in children and adolescents with 22q11.2 deletion syndrome. *Journal of Intellectual Disability Research*, *59*(9), 803–815.

Hidding, E., Swaab, H., de Sonneville, L. M. J., van Engeland, H., & Vorstman, J. A. S. (2016). The role of COMT and plasma proline in the variable penetrance of autistic spectrum symptoms in 22q11.2 deletion syndrome. *Clinical Genetics*, *90*(5), 420–427.

Higgins, J. P. T., Thompson, S. G., Deeks, J. J., & Altman, D. G. (2003). Measuring inconsistency in meta-analyses. *BMJ*, *327*(7414), 557–560.

Hoffmire, C. A., Magyar, C. I., Connolly, H. V., Fernandez, I. D., & van Wijngaarden, E. (2014). High prevalence of sleep disorders and associated comorbidities in a community sample of children with Down syndrome. *Journal of Clinical Sleep Medicine, 10*(4), 411–419.

Howlin, P., Karpf, J., & Turk, J. (2005). Behavioural characteristics and autistic features in individuals with Cohen syndrome. *European Child & Adolescent Psychiatry, 14*(2), 57–64.

Huang, C-H., Peng, S. S-F., Weng, W-C., Su, Y-N., Lee, W-T., & National Taiwan University Hospital Tuberous Sclerosis Complex (NTUH TSC) Study Group. (2015). The relationship of neuroimaging findings and neuropsychiatric comorbidities in children with tuberous sclerosis complex. *Journal of the Formosan Medical Association, 114*(9), 849–854.

Jalbrzikowski, M., Lazaro, M. T., Gao, F., Huang, A., Chow, C., Geschwind, D. H., …& Bearden, C. E. (2015). Transcriptome profiling of peripheral blood in 22q11.2 deletion syndrome reveals functional pathways related to psychosis and autism spectrum disorder. *PloS One, 10*(7), e0132542. doi: 10.1371/journal.pone.0132542

Jeste, S. S., Varcin, K. J., Hellemann, G. S., Gulsrud, A. C., Bhatt, R., Kasari, C., …& Nelson, C. A. (2016). Symptom profiles of autism spectrum disorder in tuberous sclerosis complex. *Neurology, 87*(8), 766–772.

Jeste, S. S., Wu, J. Y., Senturk, D., Varcin, K., Ko, J., McCarthy, B., …& Nelson, C. A. (2014). Early developmental trajectories associated with ASD in infants with tuberous sclerosis complex. *Neurology, 83*(2), 160–168.

Kaufmann, W. E., Kidd, S. A., Andrews, H. F., Budimirovic, D. B., Esler, A., Haas-Givler, B., …& Berry-Kravis, E. (2017). Autism spectrum disorder in Fragile X syndrome:

Cooccurring conditions and current treatment. *Pediatrics, 139*(Supplement 3), S194–S206.

Khemir, S., Halayem, S., Azzouz, H., Siala, H., Ferchichi, M., Guedria, A., …& Kaabachi, N. (2016). Autism in Phenylketonuria patients: From clinical presentation to molecular defects. *Journal of Child Neurology, 31*(7), 843–849.

Kim, S. H., & Lord, C. (2012). New Autism Diagnostic Interview-Revised algorithms for toddlers and young preschoolers from 12 to 47 months of age. *Journal of Autism and Developmental Disorders, 42*(1), 82–93.

Klusek, J., Hunt, A. W., Mirrett, P. L., Hatton, D. D., Hooper, S. R., Roberts, J. E., & Bailey, D. B. (2015). Reading and phonological skills in boys with Fragile X syndrome. *Journal of Autism and Developmental Disorders, 45*(6), 1699–1711.

Kothare, S. V., Singh, K., Hochman, T., Chalifoux, J. R., Staley, B. A., Weiner, H. L., …& Devinsky, O. (2014). Genotype/phenotype in tuberous sclerosis complex: Associations with clinical and radiologic manifestations. *Epilepsia, 55*(7), 1020–1024.

Lane, C., Milne, E., & Freeth, M. (2017). Characteristics of autism spectrum disorder in Sotos syndrome. *Journal of Autism and Developmental Disorders, 47*(1), 135–143.

Lee, M., Martin, G. E., Berry-Kravis, E., & Losh, M. (2016). A developmental, longitudinal investigation of autism phenotypic profiles in Fragile X syndrome. *Journal of Neurodevelopmental Disorders, 8*. doi: 10.1186/s11689-016-9179-0

Lipsey, M. W., & Wilson, D. B. (2001). The way in which intervention studies have 'personality' and why it is important to meta-analysis. *Evaluation & the Health Professions, 24*(3), 236–254.

Loesch, D. Z., Bui, Q. M., Dissanayake, C., Clifford, S., Gould, E., Bulhak-Paterson, D., …& Huggins, R. M. (2007). Molecular and cognitive predictors of the continuum of autistic behaviours in Fragile X. *Neuroscience and Biobehavioral Reviews*, *31*(3), 315–326.

Lord, C., Rutter, M., DiLavore, P. C., Risi, S., Gotham, K., & Bishop, S. L. (2012). *Autism Diagnostic Observation Schedule-2*. Los Angeles: Western Psychological Services.

Lord, C., Rutter, M., Goode, S., Heemsbergen, J., Jordan, H., Mawhood, L., & Schopler, E. (1989). Autism Diagnostic Observation Schedule: A standardized observation of communicative and social behavior. *Journal of Autism and Developmental Disorders*, *19*(2), 185–212.

Lord, C., Rutter, M., & Le Couteur, A. (1994). Autism Diagnostic Interview-Revised: A revised version of a diagnostic interview for caregivers of individuals with possible pervasive developmental disorders. *Journal of Autism and Developmental Disorders*, *24*(5), 659–685.

Mason, J., & Scior, K. (2004). 'Diagnostic overshadowing' amongst clinicians working with people with intellectual disabilities in the UK. *Journal of Applied Research in Intellectual Disabilities*, *17*(2), 85–90.

Moher, D., Liberati, A., Tetzlaff, J., Altman, D. G., & The PRISMA Group. (2009). Preferred Reporting Items for Systematic Reviews and Meta-Analyses: The PRISMA Statement. *PLOS Medicine*, *6*(7), e1000097. doi: 10.1371/journal.pmed.1000097

Molloy, C. A., Murray, D. S., Kinsman, A., Castillo, H., Mitchell, T., Hickey, F. J., & Patterson, B. (2009). Differences in the clinical presentation of Trisomy 21 with and without autism. *Journal of Intellectual Disability Research*, *53*(2), 143–151.

Morris, S. M., Acosta, M. T., Garg, S., Green, J., Huson, S., Legius, E., …& Constantino, J. N. (2016). Disease burden and symptom structure of autism in Neurofibromatosis type 1: A study of the International NF1-ASD Consortium Team (INFACT). *JAMA Psychiatry*, *73*(12), 1276–1284.

Moss, J., & Howlin, P. (2009). Autism spectrum disorders in genetic syndromes: Implications for diagnosis, intervention and understanding the wider autism spectrum disorder population. *Journal of Intellectual Disability Research*, *53*(10), 852–873.

Moss, J., Oliver, C., Nelson, L., Richards, C., & Hall, S. (2013). Delineating the profile of autism spectrum disorder characteristics in Cornelia de Lange and Fragile X syndromes. *American Journal on Intellectual and Developmental Disabilities*, *118*(1), 55–73.

Naerland, T., Bakke, K. A., Storvik, S., Warner, G., & Howlin, P. (2017). Age and gender-related differences in emotional and behavioural problems and autistic features in children and adolescents with Down syndrome: A survey-based study of 674 individuals. *Journal of Intellectual Disability Research*, *61*(6), 594–603.

Nomura, Y., & Segawa, M. (2005). Natural history of Rett syndrome. *Journal of Child Neurology*, *20*(9), 764–768.

Oliver, C., Berg, K., Moss, J., Arron, K., & Burbidge, C. (2011). Delineation of behavioral phenotypes in genetic syndromes: Characteristics of autism spectrum disorder, affect and hyperactivity. *Journal of Autism and Developmental Disorders*, *41*(8), 1019–1032.

Olszewski, A. K., Kikinis, Z., Gonzalez, C. S., Coman, I. L., Makris, N., Gong, X., …& Kates, W. R. (2017). The social brain network in 22q11.2 deletion syndrome: A

diffusion tensor imaging study. *Behavioral and Brain Functions*, *13*(1), 4.

Oxelgren, U. W., Myrelid, Å., Annerén, G., Ekstam, B., Göransson, C., Holmbom, A., …& Fernell, E. (2017). Prevalence of autism and attention-deficit-hyperactivity disorder in Down syndrome: A population-based study. *Developmental Medicine and Child Neurology*, *59*(3), 276–283.

Persico, A. M., & Bourgeron, T. (2006). Searching for ways out of the autism maze: Genetic, epigenetic and environmental clues. *Trends in Neurosciences*, *29*(7), 349–358.

Plasschaert, E., Descheemaeker, M., Van Eylen, L., Noens, I., Steyaert, J., & Legius, E. (2015). Prevalence of autism spectrum disorder symptoms in children with Neurofibromatosis type 1. *American Journal of Medical Genetics. Part B, Neuropsychiatric Genetics*, *168B*(1), 72–80.

Plasschaert, E., Van Eylen, L., Descheemaeker, M.-J., Noens, I., Legius, E., & Steyaert, J. (2016). Executive functioning deficits in children with Neurofibromatosis type 1: The influence of intellectual and social functioning. *American Journal of Medical Genetics. Part B, Neuropsychiatric Genetics*, *171B*(3), 348–362.

Prather, P., & de Vries, P. J. (2004). Behavioral and cognitive aspects of tuberous sclerosis complex. *Journal of Child Neurology*, *19*(9), 666–674.

Pretto, D., Yrigollen, C. M., Tang, H., Williamson, J., Espinal, G., Iwahashi, C. K., …& Tassone, F. (2014). Clinical and molecular implications of mosaicism in FMR1 full mutations. *Frontiers in Genetics*, *5*, 318.

Reilly, C., Senior, J., & Murtagh, L. (2015). ASD, ADHD, mental health conditions and psychopharmacology in neurogenetic syndromes: Parent survey. *Journal of Intellectual Disability Research*, *59*(4), 307–318.

Richards, C., Jones, C., Groves, L., Moss, J., & Oliver, C. (2015). Prevalence of autism spectrum disorder phenomenology in genetic disorders: A systematic review and meta-analysis. *The Lancet Psychiatry*, *2*(10), 909–916.

Roberts, J. E., Tonnsen, B. L., McCary, L. M., Caravella, K. E., & Shinkareva, S. V. (2016). Brief report: Autism symptoms in infants with Fragile X syndrome. *Journal of Autism and Developmental Disorders*, *46*(12), 3830–3837.

Rutter, M., Bailey, A., & Lord, C. (2003). *Social Communication Questionnaire (SCQ)*. Los Angeles: Western Psychological Services.

Sappok, T., Diefenbacher, A., Budczies, J., Schade, C., Grubich, C., Bergmann, T., …& Dziobek, I. (2013). Diagnosing autism in a clinical sample of adults with intellectual disabilities: How useful are the ADOS and the ADI-R? *Research in Developmental Disabilities*, *34*(5), 1642–1655.

Schopler, E., Van Bourgondien, M., Wellman, G., & Love, S. (2010). *Childhood Autism Rating Scale-2nd Edition*. Los Angeles: Western Psychological Services.

Schwarzer, G. (2007). Meta: An R package for meta-analysis. *R News*, *7*(3), 40–45.

Seltzer, M. M., Shattuck, P., Abbeduto, L., & Greenberg, J. S. (2004). Trajectory of development in adolescents and adults with autism. *Mental Retardation and Developmental Disabilities Research Reviews*, *10*(4), 234–247.

Senn, S. (2007). Trying to be precise about vagueness. *Statistics in Medicine*, *26*(7), 1417–1430.

Shehata, H. S., AbdelGhaffar, H. M., Nasreldin, M., Elmazny, A., Abdelalim, A., Sabbah, A., & Shalaby, N. M. (2017). Clinical patterns and outcomes of status epilepticus in patients with tuberous sclerosis complex. *Therapeutics and Clinical Risk Management*, *13*, 779–785.

Sheth, K., Moss, J., Hyland, S., Stinton, C., Cole, T., & Oliver, C. (2015). The behavioral characteristics of Sotos syndrome. *American Journal of Medical Genetics. Part A*, *167A*(12), 2945–2956.

Skuse, D. H. (2007). Rethinking the nature of genetic vulnerability to autistic spectrum disorders. *Trends in Genetics*, *23*(8), 387–395.

Talkowski, M. E., Minikel, E. V., & Gusella, J. F. (2014). Autism spectrum disorder genetics: Diverse genes with diverse clinical outcomes. *Harvard Review of Psychiatry*, *22*(2), 65–75.

Taquet, M., Scherrer, B., Commowick, O., Peters, J. M., Sahin, M., Macq, B., & Warfield, S. K. (2014). A mathematical framework for the registration and analysis of multi-fascicle models for population studies of the brain microstructure. *IEEE Transactions on Medical Imaging*, *33*(2), 504–517.

Tassone, F., Iong, K. P., Tong, T., Lo, J., Gane, L. W., Berry-Kravis, E., …& Hagerman, R. J. (2012). FMR1 CGG allele size and prevalence ascertained through newborn screening in the United States. *Genome Medicine*, *4*(12), 100.

Thurman, A. J., McDuffie, A., Kover, S. T., Hagerman, R. J., & Abbeduto, L. (2015). Autism symptomatology in boys with Fragile X syndrome: A cross sectional developmental trajectories comparison with nonsyndromic autism spectrum disorder. *Journal of Autism and Developmental Disorders*, *45*(9), 2816–2832.

Trillingsgaard, A., & Østergaard, J. R. (2004). Autism in Angelman syndrome: An exploration of comorbidity. *Autism: The International Journal of Research and Practice*, *8*(2), 163–174.

Tye, C., Farroni, T., Volein, Á., Mercure, E., Tucker, L., Johnson, M. H., & Bolton, P. F. (2015). Autism diagnosis differentiates neurophysiological responses to faces in

adults with tuberous sclerosis complex. *Journal of Neurodevelopmental Disorders, 7*, 33.

Vangkilde, A., Jepsen, J. R. M., Schmock, H., Olesen, C., Arnarsdóttir, S., Baaré, W. F. C., …& Olsen, L. (2016). Associations between social cognition, skills, and function and subclinical negative and positive symptoms in 22q11.2 deletion syndrome. *Journal of Neurodevelopmental Disorders*, *8*, 42.

Viechtbauer, W. (2010). Conducting meta-analyses in R with the metafor package. *Journal of Statistical Software, 36*(3), 1–48.

Warner, G., Moss, J., Smith, P., & Howlin, P. (2014). Autism characteristics and behavioural disturbances in ~ 500 children with Down's syndrome in England and Wales. *Autism Research, 7*(4), 433–441.

Wheeler, A. C., Mussey, J., Villagomez, A., Bishop, E., Raspa, M., Edwards, A., …& Bailey, D. B. (2015). DSM-5 changes and the prevalence of parent-reported autism spectrum symptoms in Fragile X syndrome. *Journal of Autism and Developmental Disorders*, *45*(3), 816–829.

Wilbur, C., Sanguansermsri, C., Chable, H., Anghelina, M., Peinhof, S., Anderson, K., …& Connolly, M. B. (2017). Manifestations of tuberous sclerosis complex: The experience of a provincial clinic. *The Canadian Journal of Neurological Sciences*, *44*(1), 35–43.

Wink, L. K., Fitzpatrick, S., Shaffer, R., Melnyk, S., Begtrup, A. H., Fox, E., …& Erickson, C. A. (2015). The neurobehavioral and molecular phenotype of Angelman syndrome. *American Journal of Medical Genetics. Part A*, *167A*(11), 2623–2628.

Woodman, A. C., Smith, L. E., Greenberg, J. S., & Mailick, M. R. (2015). Change in autism symptoms and maladaptive behaviors in adolescence and adulthood: The role of positive family processes. *Journal of Autism and Developmental Disorders*, *45*(1), 111–126.

World Health Organisation (1992). *ICD-10 Classifications of Mental and Behavioural Disorder: Clinical Descriptions and Diagnostic Guidelines*. Geneva: World Health Organisation.

Wulffaert, J., van Berckelaer-Onnes, I. A., & Scholte, E. M. (2009). Autistic disorder symptoms in Rett syndrome. *Autism: The International Journal of Research and Practice*, *13*(6), 567–581.

Yang, G., Shi, Z. N., Meng, Y., Shi, X. Y., Pang, L. Y., Ma, S. F., …& Zou, L. P. (2017). Phenotypic and genotypic characterization of Chinese children diagnosed with tuberous sclerosis complex. *Clinical Genetics*, *91*(5), 764–768.

Zhao, X., Anthony, L., Vlad, K., Clara, L., Daniel, G., Kiely, L., …& Michael, W. (2007). A unified genetic theory for sporadic and inherited autism. *Proceedings of the National Academy of Sciences*, *104*(31), 12831–12836.

CHAPTER 2

A developmental scale of early social cognition in autism spectrum disorder

1. Abstract

Background: 'Theory of Mind' (ToM) refers to the ability to represent the mental states of others and plays a critical role in successful social interaction. Although it was previously thought that ToM emerged around 4-5 years of age, recent evidence suggests these abilities may be present in much younger infants. Using a novel battery of tasks, Powis (2014) evidenced an understanding of both *intention* and *shared intentionality* in typically developing children aged 14-34 months. Furthermore, these skills emerged in a reliable developmental progression across the sample. Evidence suggests that children with autism spectrum disorder (ASD) are delayed in the development of more complex ToM abilities such as understanding *belief* and *desire*, however, the development of *early* social cognitive skills in children with ASD is less understood.

Method: The Early Social Cognition Scale, a battery of six tasks assessing an understanding of others' *intentions* and the ability to engage in *shared intentionality*, was administered to 21 children with ASD (mean age 8.5 years; 85.7% male). Participants were also administered the ADOS and the Mullen Scales of Early Learning.

Results: Guttman scaling analysis produced a reliable developmental scale of early social cognitive skills in children with ASD, paralleling that observed in a normative sample of typically developing children. Further analysis, however, suggested that skill acquisition was delayed in children with ASD, and that interpreting the intentions of others through eye gaze was a particularly challenging development step. Overall scale performance was positively correlated with verbal, but not non-verbal, mental age.

Discussion: Results are discussed in relation to our understanding of typical and atypical ToM development. Clinical implications for assessment and intervention in children with ASD are also considered.

2. Introduction

'Theory of Mind' (ToM), a term coined by Premack and Woodruff (1978), refers to the ability to represent the mental states of others. ToM is hypothesised to play a crucial role in successful interaction, allowing individuals to take the perspective of others and engage in complex social behaviours (Baron-Cohen, Tager-Flusberg, & Lombardo, 2013). Evidence suggests that ToM is impaired in individuals with autism spectrum disorder (ASD), and that these ToM deficits might *underpin* difficulties in social interaction and communication (Baron-Cohen, 1995; Baron-Cohen, Leslie, & Frith, 1985; Kimhi, 2014). However, ToM development in ASD is not fully understood. In particular, it is not clear whether ToM acquisition is *delayed* or *different*. This study aimed to address this question by exploring the developmental trajectory of early social cognition in children with ASD.

2.1 Theory of Mind development

ToM has traditionally been studied using *false belief* tasks, which require an individual to recognise that another person can hold and act on a belief that the individual knows to be untrue (Apperly, 2012). Despite variation in the content and format of false belief tasks, a meta-analysis of 178 studies revealed a consistent picture; whilst typically developing children usually fail tests of false belief under the age of three years, they begin to reliably pass these tasks by four to five years of age (Wellman, Cross, & Watson, 2001). Although some have argued that success on false belief tasks represents developments in skills such as language or executive functioning (see Wellman et al., 2001), the authors interpreted the consistency of their findings as evidence of a 'conceptual change' in children's understanding of others during the preschool years.

Despite this traditional focus on false belief understanding, ToM is increasingly understood to encompass a range of mentalising constructs, achieved at different developmental stages (Carruthers, 2013; Wellman & Liu, 2004). A meta-analysis by Wellman and Liu (2004), for example, revealed that children aged 3-5 years were significantly more likely to pass tests of *diverse desire* (i.e. an understanding that others have desires that are different from one's own) and *diverse belief* (i.e. an understanding that others have beliefs that are different from one's own) than tasks assessing false belief understanding. Children also performed significantly better on tests requiring a judgment of another's knowledge/ignorance than they did on tests of false belief. These findings indicate the presence of multiple ToM constructs involving different mental state judgments, and suggest that such constructs emerge at different stages of development.

2.2 A Theory of Mind 'scale'

The group level effects reported by Wellman and Liu (2004) do not establish unequivocally whether children achieve different ToM constructs in a consistent developmental sequence. It is possible, for example, that whilst some children develop an understanding of desire prior to an understanding of belief, for other children these concepts are mastered in the reverse order. Wellman and Liu (2004) explored whether a ToM 'scale', consisting of a range of ToM constructs, could yield evidence of a reliable developmental progression. Wellman and Liu (2004) recruited children aged three to five years and administered seven tasks assessing understanding of *diverse desire, diverse belief, false belief, knowledge access* (i.e. that not seeing means not knowing), and *hidden emotion* (i.e. that individuals may display an emotion discrepant with the emotion they are experiencing). Guttman scaling analyses (Guttman, 1950) revealed that typically

developing children pass these tasks in a consistent sequence, such that *diverse desire* is understood before *diverse belief*, followed by an understanding of *knowledge access*, then *false belief*, and lastly an understanding of *hidden emotion*. Wellman and Liu (2004) suggest that skills might develop through modification, such that 'easier' skills are generalised to include later-developing skills, or mediation, such that 'easier' skills form the basis for later-developing abilities. This ToM scale has been replicated in typically developing children from both the US (Wellman, Lopez-Duran, LaBounty, & Hamilton, 2008) and Australia (Peterson, Wellman, & Liu, 2005).

2.3 Autism spectrum disorder as a model of atypical Theory of Mind

Wellman, Fuxi and Peterson (2011) argue that a ToM scale could advance understanding of both typical and atypical ToM acquisition. Autism spectrum disorder (ASD), for example, is characterised by impaired social interaction and communication and restricted and repetitive interests and behaviours, estimated to occur in approximately 1% of the population (American Psychiatric Association, 2013; Baird et al., 2006). In a seminal study, Baron-Cohen et al. (1985) demonstrated that 80% of individuals with ASD aged 6-16 years failed a test of false belief, prompting hypotheses that deficits in ToM *underpin* social interaction difficulties in ASD (Baron-Cohen, 1995; Kimhi, 2014). However, studies focusing solely on false belief offer no insight into whether individuals with ASD demonstrate other forms of mentalising, or how those who do pass false belief tasks come to achieve this skill. Thus it is not clear whether ToM development is *delayed* or *different* in individuals with ASD. This distinction has implications for understanding the developmental processes through which ToM abilities emerge in both typical and atypical

populations, as well as for considering the links between ToM and social interaction in ASD.

To delineate ToM acquisition in ASD, Peterson et al. (2005) administered tasks from Wellman and Liu's (2004) scale in a sample of 36 individuals with ASD aged 6-14 years. Results indicated that these older children and adolescents showed a delay in their understanding of a range of mental state constructs. Furthermore, individuals with ASD displayed a different pattern of skill acquisition; whilst in typically developing children an understanding of *false belief* preceded an understanding of *hidden emotion*, children with ASD tended to pass a task assessing their understanding of *hidden emotion* prior to passing *false belief* tasks. It is possible, therefore, that individuals with ASD are using different strategies in their acquisition of particular ToM skills compared to typically developing children (Peterson et al., 2005). Furthermore, this finding highlights that children who display impaired ToM as assessed by traditional false belief tasks might nonetheless possess the capacity to understand other mental state constructs.

2.4 Precursors to Theory of Mind in infants

Although children younger than three years have typically failed 'explicit' ToM tasks involving overt instruction and questioning, evidence suggests that false belief understanding might, in fact, be expressed 'implicitly' (i.e. in an individual's spontaneous reactions) in infants as young as 15 months old. Southgate, Senju and Csibra (2007) found that when 25-month-old children anticipated an actor looking for an object, they looked towards the location where the actor had last seen the object, despite the child knowing the object had been moved in the actor's absence. Evidence of implicit false belief

understanding in infants has been reported by studies applying a range of paradigms, including those measuring looking-time as an indicator of infants' expectations about another's behaviour and scenarios based on beliefs about the properties of an object (i.e. 'non-search' paradigms; Kovács, Téglás, & Endress, 2010; Scott, Baillargeon, Song, & Leslie, 2010; Surian, Caldi, & Sperber, 2007). These studies suggest that typically developing children younger than three years may be able to represent the mental states of others but might fail explicit tasks due to the demands they place on cognitive processing systems (Baillargeon, Scott, & He, 2010; Carruthers, 2013).

In a separate line of research, studies have been exploring infant social cognition through the early explicit social cognitive skills that might act as a *foundation* for later-developing ToM skills. Tomasello, Carpenter, Call, Behne and Moll (2005), for example, have argued that an understanding of others' *intentions* is a precursor to later-developing, more complex ToM constructs such as understanding others' *belief*. Evidence suggests that children aged 14 months are able to interpret the communicative intention of both an adult's pointing gesture and eye gaze in order to direct their attention to a hidden object (Behne, Carpenter, & Tomasello, 2005). Furthermore, infants begin to show instrumental helping behaviour, for example passing an out-of-reach object to an adult, between 14 and 18 months of age (Dunfield, Kuhlmeier, O'Connell, & Kelley, 2011; Svetlova, Nichols, & Brownell, 2010; Warneken & Tomasello, 2006; Warneken & Tomasello, 2007). In contrast to 'helping' conditions, Warneken and Tomasello (2006) found that infants did not assist the adult in control conditions in which the behaviour was the same (e.g. dropping an object on the floor) but help was not needed (e.g. the adult did not attempt to reach for the object). Warneken and Tomasello (2006) interpreted this as evidence that infants

understood the adult's intention to obtain the object in the helping condition and acted to facilitate the adult in meeting this goal.

Meltzoff (1995) further demonstrated that children aged 18 months who viewed an actor attempting to carry out a task but 'failing' to complete it understood the actor's intention and replicated the actor's *intended* act. Consistent with Tomasello et al.'s (2005) assertion that these early social cognitive skills might form the foundation for later ToM skills, Olineck and Poulin-Dubois (2005) found that infants' ability to discriminate between accidental and intentional actions at 14 and 18 months of age was predictive of their later use of internal state language at 30 months. Colonnesi, Rieffe, Koops and Perucchini (2008) further found that the performance of children aged 12 and 15 months on a task requiring them to re-enact intended (but not performed) acts was significantly related to their ability to understand their own, prior, false belief when they were 39 months old.

Tomasello et al. (2005) suggest that, over the second year of life, infants build on this understanding of others' intentions and begin to engage in collaborative activities requiring *shared intentionality*. Tomasello et al. (2005) define shared intentionality as the creation of a joint goal with another, and the coordination of intentions and actions in pursuit of that goal. Warneken, Chen and Tomasello (2006) found that children aged 18 and 24 months old (but not 14-month-olds) engaged collaboratively with an adult in cooperative problem solving tasks and cooperative social games. Of particular note, when the adult deliberately ceased their role in the task, children aged 18 and 24 months made attempts to reengage the adult. Warneken et al. (2006) interpreted this as evidence that around 18 months of age infants begin to coordinate their own intentions and actions with the intentions of others to achieve a shared goal.

2.5 An Early Social Cognition Scale

Applying Wellman and Liu's (2004) scaling approach, Powis (2014) sought to develop a reliable scale sequencing the development of early social cognitive skills in infants. From a review of the literature, seven tasks assessing early social cognitive skills were generated, including: 1) instrumental helping, 2) understanding the communicative intent of a pointing gesture, 3) understanding the communicative intent of eye gaze, 4) re-enacting an intended but unsuccessful act, 5) understanding that seeing leads to knowing, 6) cooperation with an adult in a problem-solving task, and 7) cooperation with an adult during a social game. Powis (2014) administered these tasks to 86 children aged 14 to 34 months and demonstrated, using Guttman scaling analyses, a reliable developmental sequence of six early 'ToM precursor' skills. This showed that children first display helping behaviour, followed by the ability to understand a pointing gesture and an understanding of another's implicit intentions and goals. These skills precede the ability to understand the communicative content of gaze, and to co-ordinate with another person in a problem-solving task, and finally the ability to cooperate with another person during a social game.

This Early Social Cognition Scale complements and expands the ToM scale produced by Wellman and Liu (2004), offering opportunities for a more comprehensive and longitudinal perspective of social cognitive development. Despite the value of Wellman and Liu's (2004) scale, the processing demands of the individual tasks limit their ability to identify, and therefore to investigate, ToM abilities in younger children and children with intellectual disability. In contrast, the tasks used in the Early Social Cognition Scale were

designed to be engaging and flexible, and they require very little expressive and receptive language, making it a valuable tool for assessing social cognition in these groups.

2.6 Early social cognition in ASD

Powis' (2014) Early Social Cognition Scale has not been applied to children with ASD. Research exploring the performance of pre-school age children with ASD on individual tasks demonstrates some understanding of others' intentions. Studies by Carpenter, Pennington and Rogers (2001) and Aldridge, Stone, Sweeney and Bower (2000), for example, found that children with ASD performed equally, or better, than control groups without ASD on tasks requiring re-enactment of intended (but unsuccessful) acts. Liebal, Colombi, Rogers, Warneken and Tomasello (2008) further demonstrated instrumental helping behaviour in children with ASD with a nonverbal mental age of at least 15 months. However, these children were significantly less likely than a comparison group with developmental delay to coordinate their actions with an adult in cooperative problem-solving tasks and games. Furthermore, when the adult partner interrupted the cooperative activity, individuals with ASD made fewer attempts to re-engage their partner. Finally, individuals with ASD have consistently been shown to demonstrate a lack of joint attention behaviours, including both initiation of joint attention through pointing or showing objects, and responding to others' attempts to engage in joint attention (Baron-Cohen, 1995; Leekam, Baron-Cohen, Perrett, Milders, & Brown, 1997). These deficits might suggest that children with ASD are impaired in their ability to interpret the communicative intentions in the gestures and eye gaze of others.

Despite these initial findings, the extent to which children with ASD display these early social cognitive skills requires further exploration. In particular, the developmental sequence in which these skills are acquired in individuals with ASD is unknown. This could have important clinical implications, potentially highlighting areas of strength or difficulty to guide assessment and intervention. Furthermore, detailing the performance of children with ASD on the Early Social Cognition Scale would have implications for understanding the acquisition of these skills across both typical and atypical populations. If children with ASD demonstrate the same developmental trajectory as observed in typically developing children, this would lend weight to an argument that earlier-developing skills form the foundation for skills that are acquired later. If, however, children with ASD show a different pattern of development, or miss certain 'steps' on the scale whilst successfully acquiring later skills, this might suggest that skills can develop independently, without reliance on earlier-developing abilities.

This study uses Powis' (2014) battery of assessments to answer the following questions:

1) Do children with ASD display these early social cognitive skills?
2) If these early social cognitive skills are present, do these skills emerge in the same, or a different, developmental sequence when compared to typically developing children?

3. Method

3.1 Participants

Children with a diagnosis of ASD were recruited primarily through a special educational needs school in the West Midlands. The school was initially contacted via telephone, following which a letter (Appendix 6) and information sheet (Appendix 7) about the study was sent. After agreeing to assist recruitment, the school was sent study information sheets and opt-in consent forms for distribution to parents/carers of children with a diagnosis of ASD (Appendix 8). The parents/carers of 23 children consented to participation in the study. Six additional participants were recruited from a participant database held by the Cerebra Centre for Neurodevelopmental Disorders (CCND). These participants were contacted initially to participate in other studies at CCND, and once enrolled given the opportunity to take part in this study.

In line with protocol described in Powis (2014), Wellman and Liu's (2004) ToM Scale was initially administered to eight participants, out of the total 29, who were reported to be verbally fluent by school staff. These data are not included here as the present study focuses on data from the Early Social Cognition Scale. Wellman and Liu's (2004) ToM Scale consists of five tasks assessing an understanding of *diverse desire*, *diverse belief*, *knowledge access*, *false belief*, and *real-apparent emotion*. If children failed two of the first three tasks of Wellman and Liu's (2004) battery, they were subsequently assessed using the Early Social Cognition Scale, and their data were included in the current study (*n*=3).

The Early Social Cognition Scale was thus administered to 24 participants. Three participants were excluded because they did not engage with the study tasks (mean age

6.30 years; age range 7.25-7.75 years). The final sample therefore consisted of 21 children, of which 85.71% were male and 14.29% were female (see Table 1 for participant data).

3.2 Tasks and Measures

3.2.1 The Autism Diagnostic Observation Schedule

The Autism Diagnostic Observation Schedule (Lord et al., 1989; Lord et al., 2012) was administered by a trained researcher to confirm ASD diagnosis (see Table 1).

3.2.2 Assessment of mental age

Mental age was assessed using either the Mullen Scales of Early Learning (MSEL; Mullen, 1995) or the British Ability Scales Second Edition (BAS-II; Elliott, Smith, & McCulloch, 1996). The MSEL (Mullen, 1995) is a standardised cognitive assessment of verbal and non-verbal skills in infants and children from birth up to a mental age of 68 months. Participants were administered the Fine Motor and Visual Reception subscales (non-verbal), and the Receptive Language and Expressive Language subscales (verbal). The BAS-II (Elliott et al., 1996) is a standardised cognitive assessment of verbal and non-verbal skills for children with a mental age between 2 years 6 months to 17 years 11 months. Participants were administered the Matrices and Quantitative Reasoning subscales (non-verbal) and the Word Definitions and Verbal Similarities subscales (verbal). The most suitable assessment was selected by researchers based on the child's chronological age, preliminary information from the school regarding the child's verbal fluency, and perceived receptive and expressive language ability during initial contact with the child.

The MSEL (Mullen, 1995) provides normative data for children with a chronological age of ≤66 months only, therefore age equivalent scores were used in the present study. For each participant mean verbal and non-verbal age equivalents were calculated. A paired-samples t-test revealed a significant difference between verbal and non-verbal age equivalents ($t(18)=2.547$, $p=.02$; see Table 1).

Table 1. Participant data: chronological age, mental age and ADOS social communication scores

	Mean	Range
Age (months)	101.52	39.96 – 171.96
Non-verbal mental age equivalent (months)	40.34	6.50 – 114.00
Verbal mental age equivalent (months)	32.84	4.00 – 92.00
ADOS social communication total score[1]	13.35	8-20

[1]ADOS administered to five participants, ADOS-2 administered to the remaining participants. ADOS scores unavailable for one participant.

3.2.3 The Early Social Cognition Scale

Participants were administered Powis' (2014) six-task Early Social Cognition Scale. This battery of assessments has been shown to form a reliable Guttman scale in typically developing children, with a co-efficient of reproducibility of .96 (values above .90 indicate a reliable scale; Green, 1956) and an index of consistency of .56 (values above .50 suggest the scale is reproducible, beyond what would be expected by chance; Green, 1956). Powis (2014) evidenced 100% inter-rater reliability for the scoring of the assessment battery in typically developing children. Tasks are described briefly below with detailed administration and scoring instructions provided in Appendix 9.

Helping

The *Helping* task contained two experimental trials: 1) participants observed the experimenter 'accidentally' drop a pen on to the floor near to the child, followed by an unsuccessful attempt to reach for it, and 2) the experimenter used a set of tongs to transfer three foam cones from the table into a box, then attempted to reach for three further cones that were placed next to the child. In order to pass the *Helping* task, the infant was required to pass the reached-for object to the experimenter in *one* of the two experimental trials.

For both trials a corresponding control trial was also administered, to ensure that children were correctly interpreting the experimenter's intention to access the object in the experimental trials. For Trial 1, participants observed the experimenter deliberately throw a pen on to the floor, with no attempt to reach for it. For Trial 2, the experimenter used the tongs to lift three cones and place them back on to the table, and did not reach for the cones placed next to the child.

Re-enactment of Intended Acts

This task consisted of three experimental trials. In each trial, the experimenter made three 'unsuccessful' attempts to perform a target act using a pair of objects. For each trial, after observing the failed attempts, the child was given the pair of objects accompanied by the words "Oh look what I have here", "What's this?" or "Now it's your turn". Participants were required to successfully reproduce two of the three target acts in order to pass the task.

Gestures: Point and Gestures: Gaze

In both the *Gestures: Point* and *Gestures: Gaze* tasks, Experimenter 1 showed the child a toy, before saying "Now I'll hide it" and placing it in one of two boxes concealed behind a movable screen. In the experimental conditions, Experimenter 2 indicated to the child that she was watching Experimenter 1 hide the toy by alternating her gaze between the child and the boxes and saying "I can see". Once the toy was hidden, the screen was removed. Whilst Experimenter 1 was turned away, Experimenter 2 provided a communicative gesture (*Point*: extending index finger toward the correct box; *Gaze*: gazing between the correct box and back to the infant) to indicate the location of the toy, along with raised eyebrows to express intent. Two experimental trials were administered for each gesture, and participants were required to correctly identify the location of the toy in both experimental trials in order to pass the task.

Corresponding control trials were designed to ensure that correct responses were not due to low-level attentional cues. In these control trials, after removing the screen, Experimenter 1 gave one of two non-communicative cues. This was either a 'distracted point' (hand held out with an extended index finger but looking down with an expression indicating preoccupation with something on the hand), or a 'control gaze' (gazing at the box with an absent minded facial expression). Experimenter 2 did not comment that they were watching whilst the object was hidden. Participants were administered two control trials for each gesture (*Point* and *Gaze*).

Cooperation: Tubes-with-handles and Cooperation: Trampoline

In the *Cooperation* tasks participants were required to: a) cooperate with an adult in pursuit of a shared goal, and b) evidence an attempt to re-engage their partner when the adult interrupted this joint activity. Experimenter 1 and 2 initially carried out a demonstration of the task. For *Cooperation: Tubes-with-handles*, each individual pulled a handle at either end of two overlapping tubes in the opposite direction (i.e. pulling away from one another), in order to release a toy contained within the inner tube. In the *Cooperation: Trampoline* task, two individuals were required to work jointly in order to bounce an object on a hand-held trampoline. For each task the demonstration was followed by four trials. In Trial 1 and Trial 2, the child was required to engage jointly with Examiner 1 as observed. In Trial 3 and Trial 4, however, after beginning the activity, Experimenter 1 then ceased performing their role, letting go of the object and looking down with their hands on the floor. Experimenter 1 held this position for 15 seconds, after which they resumed their role as before. During this 'interruption' period, the child's actions were coded by Experimenter 2 for attempts to re-engage Experimenter 1 in the task.

Coding schema for the *Cooperation: Tubes-with-handles* and *Cooperation: Trampoline* tasks are provided in Tables 2 and 3. In order to pass the task children were required to score a median of three for Coordination/Engagement, and to display at least one attempt at re-engagement during the interruption period.

Table 2. Coding schema for performance on the *Cooperation: Tubes-with-handles* task

Coordination	
Category	***Definition***
No success (Score=0)	Tubes not opened.
Uncoordinated (Score=1)	Success after more than 5 seconds of inappropriate actions such as standing on wrong side, letting tube drop more than once, individual play, or individual attempts.
Coordinated (Score=2)	Success, but some inappropriate actions, but not for more than 5 seconds; releasing handle not more than once.
Very coordinated (Score=3)	Success after immediate understanding of their role. Infant positions herself in correct location and performs the correct action without mistakes.
Behaviour during interruption period	
Category	***Definition***
Disengagement	Infant leaves apparatus or plays without pursuing the goal by banging the apparatus, climbing on it, etc.
Individual attempt	Infant attempts to retrieve the object individually (infant attempts to hold both handles or peel it open on one side) or attempts to continue the game alone.
Waiting	Infant remains on correct side of the apparatus, ready to perform their role.
Re-engagement	Infant is ready to perform their role and in addition tries to re-engage E1, e.g. pushing the tube, pointing at the object and vocalising whilst looking at the partner.

Table 3. Coding schema for the *Cooperation: Trampoline* task

Engagement	
Category	***Definition***
No success (Score=0)	Infant does not hold and lift trampoline.
Low engagement (Score=1)	Joint play but lots of stopping and not too excited. Infant needs a lot of persuasion.
Medium engagement (Score=2)	Some stopping or not too excited.
High engagement (Score=3)	Continuous play and rather excited (placing block on trampoline; initiating play; active shaking).
Behaviour during interruption period	
Category	***Definition***
Disengagement	Infant leaves apparatus or plays without pursuing the goal by banging the apparatus, climbing on it, etc.
Individual attempt	Infant attempts to retrieve the object individually (infant attempts to hold both handles or peel it open on one side) or attempts to continue the game alone.
Waiting	Infant remains on correct side of the apparatus, ready to perform their role.
Re-engagement	Infant is ready to perform their role and in addition tries to re-engage E1, e.g. pushing the tube, pointing at the object and vocalising whilst looking at the partner.

3.3 Procedure

Ethical approval for the study was obtained from the Science, Technology, Engineering and Mathematics Ethical Review Committee at the University of Birmingham (Appendix 10). Participants were tested individually in a quiet room at their school. Participants that were recruited through the CCND participant database were tested individually in a quiet room at the University of Birmingham. All assessments were completed by researchers trained in administration and scoring. Participants that were tested at school were typically

assessed across two sessions, to ensure they did not become fatigued. Participants generally completed the ADOS and MSEL or BAS-II in session one, and the Early Social Cognition Scale (or ToM Scale) in session two, however, some flexibility was required to accommodate children's availability during the school timetable. Participants who were tested at the University of Birmingham conducted all assessments on the same day separated by breaks between assessments as needed. The Early Social Cognition Scale was administered in one of six counterbalanced orders (detailed in Table 4).

Table 4. Task orders for administration of the Early Social Cognition Scale

	Order 1	**Order 2**	**Order 3**	**Order 4**	**Order 5**	**Order 6**
Task 1	Re-enactment of Intended Acts	Helping (control trials)	Helping (control trials)	Re-enactment of Intended Acts	Gestures (Point and Gaze)	Gestures (Point and Gaze)
Task 2	Helping (control trials)	Re-enactment of Intended Acts	Gestures (Point and Gaze)	Gestures (Point and Gaze)	Re-enactment of Intended Acts	Helping (control trials)
Task 3	Gestures (Point and Gaze)	Gestures (Point and Gaze)	Re-enactment of Intended Acts	Helping (control trials)	Helping (control trials)	Re-enactment of Intended Acts
Task 4	Cooperation – Tubes	Cooperation – Trampoline	Helping (experimental trials)	Cooperation – Tubes	Cooperation - Trampoline	Helping (experimental trials)
Task 5	Cooperation – Trampoline	Cooperation – Tubes	Cooperation – Trampoline	Helping (experimental trials)	Helping (experimental trials)	Cooperation – Tubes
Task 6	Helping (experimental trials)	Helping (experimental trials)	Cooperation – Tubes	Cooperation – Trampoline	Cooperation – Tubes	Cooperation – Trampoline

3.4 Data analysis

Guttman scaling analysis (Guttman, 1950) was used to explore whether children with ASD demonstrated a reliable developmental progression in early social cognitive skills, as assessed by the Early Social Cognition Scale. In determining whether tasks are scalable,

scalogram analyses take two factors into account. The first, the co-efficient of reproducibility, assesses the extent to which the sequence of task passes/fails diverges from a 'perfect' scale (i.e. in a perfect scale, after failing one task a child would fail all subsequent tasks). According to Green's (1956) method, a co-efficient of reproducibility ≥.90 indicates a reproducible scale. The second, the index of consistency, assesses whether the co-efficient of reproducibility is above what could be expected by chance. Green (1956) suggests an index of consistency ≥.50.

4. Preliminary analyses

Preliminary analyses of the *Helping*, *Gestures: Point* and *Gestures: Gaze* control trials were conducted to ensure that successful performance reflected an understanding of *intention*, rather than a response to low-level attentional cues.

4.1 *Helping* control trials

The *Helping* task control trials involved scenarios in which the examiner's behaviour was matched to that observed in the experimental trials but 'helping' behaviour was not required. Only 1 (4.74%) child showed helping behaviour during a control trial. This indicated that the helping behaviour observed in experimental trials reflected an understanding of the examiner's *intention* to access the reached-for object, and a motivation to help the examiner with their unmet goal.

4.2 *Gestures*: *Point* and *Gestures: Gaze* control trials

Control trials in the *Gestures* tasks consisted of non-communicative cues (namely a 'distracted point' and a 'control gaze'), designed to ensure correct responses were not the result of low-level attentional cueing. For each task (*Point* and *Gaze*), Wilcoxon signed-rank tests were used to analyse the control trial performance of those who had passed the task (i.e. those who had successfully reached for the object in both experimental trials). Analyses revealed no significant difference between the number of correct and incorrect responses for either the *Gestures: Point* control trials (Z=-1.633, p=.102), or the *Gestures: Gaze* control trials (Z=-1.000, p=.317). These analyses indicate that the control gestures had not been sufficient to direct participants' attention to the location of the toy, and

therefore that successful task completion occurred when children understood the *intention* behind the examiner's communicative gesture during the experimental trials.

5. Results

5.1 Overall performance on the Early Social Cognition Scale

Table 5 presents the pass rate for each task in the Early Social Cognition Scale, alongside pass rates from Powis' (2014) normative data for typically developing children (mean age=22 months; age range=14-34 months). These data indicate that, at a group level, the performance of children with ASD followed a similar sequence to that observed in typically developing infants. The pass rate for the *Gestures: Gaze* task (19.05%) was much lower than the *Gestures: Point* and *Re-enactment of Intended Acts* tasks (both 61.90%) in the current sample, suggesting children with ASD had particular difficulty interpreting intention in eye gaze. Across all tasks a smaller proportion of the current sample was successful compared to typically developing children. Furthermore, no children with ASD passed the *Cooperation*: *Trampoline* task, compared to 22% of typically developing children.

Table 5. Early Social Cognition Scale task pass rates in descending order

Task	**Pass rate**	
	Children with ASD (current sample)	*Typically developing children (from Powis, 2014)*
Helping	76.19%	*88%*
Gestures: Point	61.90%	*67%*
Re-enactment of intended acts	61.90%	*63%*
Gestures: Gaze	19.05%	*43%*
Cooperation: Tubes	9.52%	*37%*
Cooperation: Trampoline	0.00%	*22%*

Table 6 presents mental age data for the current sample alongside chronological age data from Powis (2014). Mental age data were used in the current study to provide a more accurate measure of ability level in children with intellectual disability. The data in Table 6 demonstrate that mean mental age in the current sample was higher than the chronological age (and therefore assumed mental age) of Powis' (2014) typically developing sample. As such, lower pass rates in the current sample suggest a delay or deficit in early social cognitive skills in ASD. Spearman's rank-order correlation analyses revealed that the number of tasks an individual passed was significantly positively correlated with verbal mental age ($r_s(15)=.549$, $p=.015$), but not non-verbal mental age ($r_s(15)=.407$, $p=.084$) in the current sample.

Table 6. Mental age data for the current sample, with chronological age data from Powis (2014)

	Children with ASD (current sample)		***Typically developing children (from Powis, 2014)***
	Non-verbal mental age equivalent	Verbal mental age equivalent	*Chronological age*
Mean (months)	40.34	32.84	*22*
Range (months)	6.50 – 114.00	4.00 – 92.00	*14 – 34*

5.2 Pairwise task comparisons

To further explore task sequencing, McNemar's tests, applying Yates' correction for continuity, were used to compare performance between adjacent tasks (ranked by pass rate). To account for the equal pass rate for *Gestures: Point* and *Re-enactment of Intended Acts* in the current sample, both of these tasks were compared to the *Helping* and the *Gestures: Gaze* tasks. McNemar's pairwise comparisons revealed that significantly more children passed the *Re-enactment of Intended Acts* task and the *Gestures: Point* task

compared to the *Gestures: Gaze* task (see Table 7), although these comparisons were not statistically significant when applying Bonferroni-Holm correction for family wise error. No other pairwise comparisons reached statistical significance.

Table 7. McNemar's pairwise comparisons, with tasks presented in descending order according to percentage pass rate

Task	Pass rate	Pairwise comparisons
Helping	76.19%	Helping - REI p=.250; Helping – Gestures: Point p=.453
Re-enactment of intended acts (REI) **Gestures: Point**	61.90%	REI – Gestures: Gaze p=.022[1] Point – Gestures: Gaze p=.012[2]
Gestures: Gaze	19.05%	Gestures: Gaze - Tubes p=.500
Cooperation: Tubes	9.52%	Tubes - Trampoline p=.500
Cooperation: Trampoline	0%	

[1] Required significance level after Bonferroni-Holm correction: $p \leq 0.010$
[2] Required significance level after Bonferroni-Holm correction: $p \leq 0.008$

5.3 Guttman scaling analyses

The above analyses indicate a sequential task progression on the Early Social Cognition Scale in children with ASD, paralleling that observed in typically developing infants (Powis, 2014). However, such comparisons do not show whether children passed these tasks in a reliable sequence, such that success on more 'difficult' tasks was achieved only by children who had passed the preceding 'easier' tasks. Guttman scaling analyses were therefore performed, ranking tasks according to the Early Social Cognition Scale (see

Figure 1). In Powis (2014), two pairs of tasks were considered to be of equal difficulty (*Re-enactment of Intended Acts* and *Gestures: Point*; and *Gestures: Gaze* and *Cooperation: Tubes*). For these task pairs, a pass is assigned if the child was successful on *either* task.

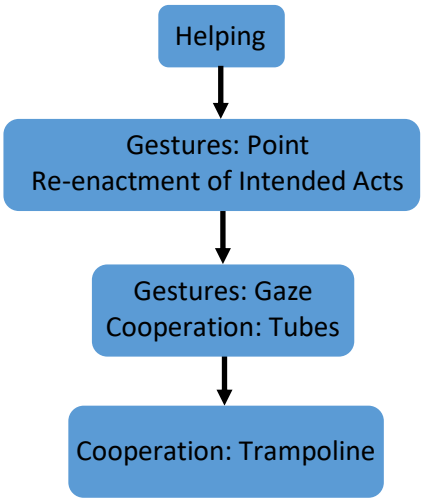

Figure 1. Early Social Cognition Scale (Powis, 2014)

In the current sample, 90.48% of participants demonstrated a pattern of performance corresponding to the Early Social Cognition Scale precisely (see Table 8). The co-efficient of reproducibility was 0.98, with an index of consistency of .58, indicating a reliable scale. Thirteen (61.90%) participants demonstrated a single pattern of performance, in which they were successful on the first two steps of the scale (Step 1: *Helping*; Step 2: *Re-enactment of intended acts* OR *Gestures: Point*) but failed all subsequent tasks.

Table 8. Group scaling according to the Early Social Cognition Scale (Powis, 2014)

Task	Pattern					
	0	**1**	**2**	**3**	**4**	**Other patterns**
Helping	−	+	+	+	+	
REI OR Gestures: Point	−	−	+	+	+	
Gestures: Gaze OR Cooperation: Tubes	−	−	−	+	+	
Cooperation: Trampoline	−	−	−	−	+	
Number of cases	3	1	13	2	0	2
Non-verbal mental age (mean/range)	29.0 *(6.5–68.5)*	13.5 *(13.5)*	40.5[1] *(17.5-62.5)*	72.8 *(31.5-114.0)*	NA	34.5[1] *(34.5)*
Verbal mental age (mean/range)	43.3 *(4.0-66.0)*	7.5 *(7.5)*	31.04 *(7.0-72.5)*	131 *(39.0-92.0)*	NA	37.0 *(37.0)*

[1] Mental age data not available for one participant

In summary, analysis of performance on the Early Social Cognition Scale revealed a reliable Guttman scale in children with ASD, paralleling the performance of a normative sample of typically developing children (Powis, 2014). However, descriptive analysis of pass rate data suggested children with ASD may be delayed in their acquisition of early social cognition skills compared to typical development. Further analyses revealed that 61.90% of the sample were successful on the first two tasks but failed *Gestures: Gaze* and both subsequent tasks, suggesting this task represented a particular challenge for children with ASD. Furthermore, no children achieved success on the final task of the scale (*Cooperation: Trampoline).* Correlational analyses showed that overall performance on the Early Social Cognition Scale (measured by the number of tasks passed) was positively associated with verbal, but not non-verbal, mental age.

6. Discussion

This study is the first to: 1) explore whether children with ASD display early explicit social cognitive skills evidencing understanding of *intention* and *shared intentionality*, as assessed by the Early Social Cognition Scale, and 2) explore the *sequence* by which children with ASD achieve success on these tasks. The study advances earlier research, which has focused on the assessment of isolated social cognitive skills, by using scaling analyses to examine the *developmental progression* of early social cognition in children with ASD. Results demonstrated a reliable developmental sequence paralleling that observed in a normative sample of typically developing infants (Powis, 2014). However, further analyses suggested possible delays, or deficits, in the acquisition of early social cognitive skills in ASD.

6.1 Early social cognitive skills in children with ASD

Results showed that children with ASD demonstrated an understanding of others' intentions in a range of tasks, including performing helping behaviour, re-enacting observed intended (but unsuccessful) acts, and interpreting communicative cues such as eye gaze and pointing gestures. These findings add to a small body of research evidencing understanding of intention in ASD equivalent to that of control groups (e.g. Aldridge et al., 2000; Carpenter et al., 2001; Liebal et al., 2008). In the current study, two (9.52%) children also demonstrated shared intentionality. In contrast to studies exploring intention, previous research reports that children with ASD are less likely than controls to engage in such shared intentionality, and make fewer attempts to engage their partner when this activity is interrupted (e.g. Liebal et al., 2008).

The current study extended these findings, which have focused on comparing group-level performance on individual tasks, by exploring the *developmental progression* of early social cognitive skills in children with ASD. Guttman scaling analyses revealed that, when tasks were ranked according to Powis' (2014) Early Social Cognition Scale, children with ASD demonstrated a reliable developmental progression consistent with that observed in typical development. Thus, children with ASD first demonstrated an understanding of another's intentions and a motivation to assist them in reaching their goals through helping behaviour. This was followed by the ability to understand the communicative intent of a pointing gesture and to understand another's implicit intentions during unsuccessful acts. These skills preceded the ability to understand the communicative content of eye gaze, to co-ordinate with another person in a problem-solving task, and finally the ability to cooperate with another person as part of a social game. Taken together with the evidence from Powis' (2014) normative sample, the results from the current study indicate that infants with ASD develop early social cognitive abilities in the *same sequence* as observed in typically developing children.

Authors adopting scaling methodology have suggested that reliable scales might indicate a process of modification, in which earlier-developing skills are generalised to include those developing later, or a process of mediation, through which 'easier' skills form a scaffold for later-developing abilities (e.g. Powis, 2014; Wellman & Liu, 2004). Although the current findings are not able to distinguish between these two explanations, the consistency between the developmental progression observed in children with ASD and in Powis' (2014) normative sample of typically developing children is in keeping with the notion that earlier skills act as a foundation for later emerging skills. Wellman et al. (2011) explored this notion further using their Theory of Mind Scale, demonstrating that performance on

the scale significantly predicted performance when re-tested a year later. Future research should use the Early Social Cognition Scale to explore the extent to which earlier performance is *predictive* of later social cognition in children with ASD.

Correlation analyses revealed that success across the scale (measured by the number of tasks passed) was significantly positively associated with verbal but not non-verbal mental age. This finding is consistent with evidence that language ability correlates with success on tasks assessing later-developing ToM abilities such as false belief (e.g. Astington & Jenkins, 1999; Milligan, Astington, & Dack, 2007). Whilst some have argued that this association is a result of the format of traditional false belief tasks, which typically require children to follow a verbal narrative and to respond to a verbally-presented question, others suggest that language might support ToM development more broadly (see Milligan et al., 2007). Linguistic development might, for example, provide labels through which children are able to represent and explore mental constructs. Furthermore, conversation with others could provide a greater awareness of differing perspectives, and therefore an understanding of others' desires, beliefs and intentions (see de Villiers & de Villiers, 2014). The Early Social Cognition Scale tasks required an understanding of very simple verbal instructions only (e.g. "Your turn"), and the examiner's actions ensured that children with limited receptive language could take part. The current findings therefore add weight to the argument that language plays a role beyond task specific factors, evidencing a connection between verbal ability and social cognitive development even when the tasks themselves require little expressive or receptive language.

6.2 Differences in early social cognition in children with ASD

Although the pattern of performance in the current sample produced a reliable scale paralleling that reported in Powis' (2014) normative data, pass rates were lower on all tasks when compared to pass rates in Powis' (2014) typically developing infants. Given that the children in the current study had a higher mean mental age, lower pass rates indicate a delay in the development of these early social cognitive skills. Thus, whilst infants with ASD might develop early social cognitive abilities in the same *sequence*, they appear to be *delayed* in their understanding of intentions and shared intentionality.

The results indicated two further distinctions in early social cognitive development in children with ASD. Firstly, percentage pass rates suggested that understanding the intentions of another through eye gaze (as assessed by the *Gestures: Gaze* task) represented a critical stage of the scale in children with ASD. Indeed, 61.90% of the sample displayed a single pattern of performance, in which they were successful on the first two steps of the scale, but then failed *Gestures: Gaze* and both subsequent tasks. It is arguable, therefore, that whilst typically developing children demonstrate a consistent sequential progression across tasks in the Early Social Cognition Scale, interpreting the eye gaze of another represents a particularly challenging transition for children with ASD.

A deficit in using eye gaze to initiate or respond to joint attention has been widely acknowledged as an early 'symptom' of ASD (Nation & Penny, 2008), and is hypothesised to be a precursor to impairments in later, more complex ToM representations (Baron-Cohen, 1995). More broadly, reduced eye contact with others is a characteristic feature of ASD, recognised within 'gold standard' diagnostic instruments such as the Autism

Diagnostic Observation Schedule (ADOS-2; Lord et al., 2012). It is possible that children with ASD do not interpret the *intentional* cue in eye gaze and therefore fail to appreciate the value of attending to another's eyes (Baron-Cohen, 1995). Alternatively, a more basic deficit in orienting to social stimuli, due to differences in attentional or reward systems, might limit an individual's opportunities for learning the relevance of eye gaze (see Nation & Penny, 2008, for a review). Nation and Penny (2008) speculate that this may be a reciprocal relationship, in which an initial lack of interest in social stimuli results in fewer opportunities for learning the significance of these social cues, leading to further deficits in orienting to eye gaze. Children with ASD may, therefore, depend on more explicit social cues, such as pointing, to direct their attention (Leekam & Ramsden, 2006). Consistent with this idea, the current findings showed that many children were able to interpret the intention behind a non-verbal communicative act when this was in the form of a pointing gesture, but did not attend to, or did not understand the intention behind, communicative eye gaze. The results are also consistent with the notion that eye gaze is integral for later-developing social cognitive skills; all those who failed the *Gestures: Gaze* task also failed the subsequent tasks of the Early Social Cognition Scale. However, these findings require replication. Again, future research could further explore the extent to which performance on this task *predicts* later social cognitive abilities.

The second distinction of note regarded children's performance on the final task of the scale, the *Cooperation: Trampoline* task. Specifically, no children in this study were successful on the task, compared to 22% of Powis' (2014) normative group. This might represent an extension of an overall delay in social cognitive development, such that none of the children reached the developmental level required for success on the final stage of the scale. Alternatively, it is possible that children with ASD demonstrated a particular

deficit in this social cognitive skill. The lack of success on the *Cooperation: Trampoline* task is of particular interest given that the task format was comparable with the *Cooperation: Tubes* task, which was passed by 9.52% of children with ASD. The crucial difference between these tasks was the *goal* of the cooperative act; for *Cooperation: Tubes* this was to retrieve a hidden toy, whereas the *Cooperation: Trampoline* task had an inherently social goal of engaging with another in a social game (Warneken et al., 2006). One possible explanation is that children with ASD were not *motivated* to engage in the *Cooperation: Trampoline* task because they experience less of a reward from such social interactions. The social motivation theory suggests that biological and psychological mechanisms bias typically developing individuals toward attending to social stimuli, experiencing reward from interactions with others, and making efforts to sustain social relationships. In contrast, these social motivational systems are hypothesised to be impaired in children with ASD (Chevallier, Kohls, Troiani, Brodkin, & Schultz, 2012). As such, some children with ASD may have possessed a capacity for shared intentionality (evidenced by their performance on the *Cooperation: Tubes* task) but lacked the motivation to engage in an inherently social act such as the *Cooperation: Trampoline* task.

6.3 Clinical implications

ToM abilities play a critical role in social interaction and deficits in ToM have been linked with a range of social and communication impairments, including difficulties engaging with others in conversation, understanding others' narrative, and responding to social or emotional cues (Kimhi, 2014). Historically, deficits in false belief understanding have been interpreted as evidence that children with ASD lack ToM (Baron-Cohen et al., 1985). The current findings, however, are consistent with the notion that social cognition encompasses

multiple mental state representations emerging at different stages of development, and that children with ASD display a range of social cognitive skills. Nonetheless, the results also highlight areas of difficulty in social cognitive development, which have implications for clinical intervention.

Findings suggested that interpreting intention in eye gaze was particularly difficult for children with ASD, and that this ability might be necessary for the acquisition of subsequent social cognitive skills. Wider evidence suggests that initiating and responding to joint attention may also be critical factors in the development of skills such as expressive language (Dawson et al., 2004). As such, approaches that support the development of joint attention might be important targets for clinical intervention. A recent review of intervention research by Green and Garg (2018) suggested that approaches based on the JASPER model (Joint Attention, Symbolic Play, Engagement and Regulation; see e.g. Kasari et al., 2014; Kasari, Gulsrud, Wong, Kwon, & Locke, 2010) bring about specific increases in joint attention behaviour in children with ASD. Furthermore, Poslawsky et al. (2015) found positive effects on children's joint attention skills following implementation of the VIPP-AUTI programme (Video-feedback Intervention to promote Positive Parenting Adapted to Autism). Further research should explore the extent to which such improvements in joint attention skills following intervention generalise to developments in other, more complex social cognitive skills.

The current results also indicated that some children with ASD had a capacity for shared intentionality, but struggled to employ these skills in the inherently social context of a cooperative game. Jahr, Eldevik and Eikeseth (2000) provided evidence for improved cooperative play in children with ASD following an intervention in which play was

supported by modelling and verbal commentary. The degree to which such 'taught' play is representative of play in typical development is the subject of some debate (see e.g. Luckett, Bundy, & Roberts, 2007), thus these findings should be considered with caution. Nonetheless, they suggest that children with ASD might be capable of developing cooperative play when provided with appropriate scaffolding.

Finally, establishing the profile of social cognition in ASD has important implications for understanding and assessing ASD in the context of other neurodevelopmental disorders. Individuals with rare genetic syndromes including Fragile X syndrome, Rett syndrome, Cohen syndrome and Cornelia de Lange syndrome are more likely to meet ASD diagnostic criteria, but appear to show distinct patterns of ASD symptomatology. It has thus been proposed that similar observable behaviours might be underpinned by different psychological mechanisms (Moss & Howlin, 2009). Cornish et al. (2007), for example, suggest that individuals with Fragile X syndrome demonstrate a good understanding of social interaction but display gaze avoidant behaviour due to social anxiety and hypersensitivity. Developing a thorough understanding of the profile of social cognition in ASD will help to differentiate the mechanisms underlying atypical social interaction in ASD and other neurodevelopmental disorders, to ensure that individuals are provided with appropriately targeted intervention.

6.4 Limitations

The above discussion should be considered in light of a number of theoretical and practical considerations. Firstly, it is important to note that a reliable Guttman scale does not, in itself, denote the progressive development of a single underlying trait or concept. Indeed, a

Guttman scale can, theoretically, be produced by measuring performance across isolated tasks, if these tasks are sufficiently varied in difficulty. Alternatively, a reliable Guttman scale in the current study could represent the development of another, unmeasured skill, such as working memory or executive functioning (Wellman et al., 2011). It is therefore of interest that performance on the Early Social Cognition Scale was positively correlated with verbal mental age in the current sample. Nonetheless, the study provides important new insights into the *sequencing* of these early social cognitive skills in children with ASD, which cannot be obtained from studies that solely compare group means on individual tasks.

A key practical consideration concerns the lack of an appropriately matched comparison group. Powis' (2014) normative data from typically developing children provided a useful benchmark from which to consider the sequencing of social cognitive development in ASD, and the current study aimed to control for differences in ability level by recruiting individuals with a mental age equivalent to the infants in Powis' (2014) normative sample. However, in order to also recruit individuals with a diagnosis of ASD, this necessarily resulted in a sample of children of an older chronological age[1] with some level of intellectual disability, which introduced other group-level differences. Jarrold and Brock (2004), for example, suggest that additional 'experience' gained with age has the potential to mask or confound group differences when making such comparisons. Furthermore, without a comparison group of children with intellectual disability *without* ASD, it is not possible to determine the specificity of the results to social cognition in ASD.

[1] NICE (2011) cautions that a diagnosis of ASD may be uncertain in children younger than 24 months of age.

Despite the limitations above, this study offers significant new contributions to the literature. By using the Early Social Cognition Scale, the study was able to assess multiple components of early social cognition in ASD, compared to previous research that has focused on a single construct such as false belief. Furthermore, this is the first study to explore the developmental progression of these early skills in children with ASD, a group at high risk of impairments in social cognition. Given the possible links between these early skills and later, more complex mentalising abilities, and the impact of ToM deficits on social interaction and communication, understanding the development of early social cognition has a key role to play in understanding the social difficulties experienced by individuals with ASD, and therefore in informing targets for intervention.

6.5 Conclusion

In conclusion, Guttman scaling revealed a reliable developmental trajectory of early social cognitive skills in children with ASD, which paralleled that previously observed in a normative sample of typically developing children. Further analyses suggested that understanding intention in the eye gaze of others represented a critical stage in the scale for children with ASD. Furthermore, no children were successful on a task requiring cooperation with another as part of a social game, indicating a deficit, or delay, in this social cognitive skill. These findings indicate areas of focus for clinical intervention, as well as areas of intact social cognitive understanding in children with ASD. Reliable scales, such as that observed here, provide some indication that earlier skills form a foundation for later-developing abilities. Future research should further these findings by examining the extent to which earlier scale performance is *predictive* of later skill development.

7. References

Aldridge, M. A., Stone, K. R., Sweeney, M. H., & Bower, T. G. R. (2000). Preverbal children with autism understand the intentions of others. *Developmental Science*, *3*(3), 294–301.

American Psychiatric Association (2013). *Diagnostic and Statistical Manual of Mental Disorders.* (5th ed.). Arlington, VA: American Psychiatric Publishing.

Apperly, I. A. (2012). What is 'theory of mind'? Concepts, cognitive processes and individual differences. *Quarterly Journal of Experimental Psychology*, *65*(5), 825–839.

Astington, J. W., & Jenkins, J. M. (1999). A longitudinal study of the relation between language and theory-of-mind development. *Developmental Psychology*, *35*(5), 1311–1320.

Baillargeon, R., Scott, R. M., & He, Z. (2010). False-belief understanding in infants. *Trends in Cognitive Sciences*, *14*(3), 110–118.

Baird, G., Simonoff, E., Pickles, A., Chandler, S., Loucas, T., Meldrum, D., & Charman, T. (2006). Prevalence of disorders of the autism spectrum in a population cohort of children in South Thames: the Special Needs and Autism Project (SNAP). *The Lancet*, *368*(9531), 210–215.

Baron-Cohen, S. (1995). *Mindblindness: An Essay on Autism and Theory of Mind.* Oxford: MIT Press.

Baron-Cohen, S., Leslie, A. M., & Frith, U. (1985). Does the autistic child have a 'theory of mind'? *Cognition*, *21*(1), 37–46.

Baron-Cohen, S., Tager-Flusberg, H., & Lombardo, M. (Eds.). (2013). *Understanding Other Minds: Perspectives from Developmental Social Neuroscience*. New York: Oxford University Press.

Behne, T., Carpenter, M., & Tomasello, M. (2005). One-year-olds comprehend the communicative intentions behind gestures in a hiding game. *Developmental Science*, *8*(6), 492–499.

Carpenter, M., Pennington, B. F., & Rogers, S. J. (2001). Understanding of others' intentions in children with autism. *Journal of Autism and Developmental Disorders*, *31*(6), 589–599.

Carruthers, P. (2013). Mindreading in infancy. *Mind & Language*, *28*(2), 141–172.

Chevallier, C., Kohls, G., Troiani, V., Brodkin, E. S., & Schultz, R. T. (2012). The social motivation theory of autism. *Trends in Cognitive Sciences*, *16*(4), 231–239.

Colonnesi, C., Rieffe, C., Koops, W., & Perucchini, P. (2008). Precursors of a theory of mind: A longitudinal study. *British Journal of Developmental Psychology*, *26*(4), 561–577.

Cornish, K., Turk, J., & Levitas, A. (2007). Fragile X Syndrome and autism: Common developmental pathways? *Current Paediatric Reviews*, *3*(1), 61-68.

Dawson, G., Toth, K., Abbott, R., Osterling, J., Munson, J., Estes, A., & Liaw, J. (2004). Early social attention impairments in autism: Social orienting, joint attention, and attention to distress. *Developmental Psychology*, *40*(2), 271–283.

de Villiers, J. G., & de Villiers, P. A. (2014). The role of language in theory of mind development. *Topics in Language Disorders*, *34*(4), 313–328.

Dunfield, K., Kuhlmeier, V. A., O'Connell, L., & Kelley, E. (2011). Examining the diversity of prosocial behavior: Helping, sharing, and comforting in infancy. *Infancy*, *16*(3), 227–247.

Elliott, C. D., Smith, P., & McCulloch, K. (1996). *British Ability Scales Second Edition (BAS II): Administration and Scoring Manual*. London: NFER-Nelson.

Green, B.F. (1956). A method of scalogram analysis using summary statistics. *Psychometrika, 21,* 79-88.

Green, J., & Garg, S. (2018). Annual Research Review: The state of autism intervention science: Progress, target psychological and biological mechanisms and future prospects. *Journal of Child Psychology and Psychiatry, and Allied Disciplines*, *59*(4), 424–443.

Guttman, L. (1950). The basis of scalogram analysis. In S.A. Stouffer, L. Guttman, E.A. Suchman, & P.A. Lazarfeld (Eds.), *Measurement and Prediction* (pp. 60-90). Princetown, NJ: Princetown University Press.

Jahr, E., Eldevik, S., & Eikeseth, S. (2000). Teaching children with autism to initiate and sustain cooperative play. *Research in Developmental Disabilities*, *21*(2), 151–169.

Jarrold, C., & Brock, J. (2004). To match or not to match? Methodological issues in autism-related research. *Journal of Autism and Developmental Disorders*, *34*(1), 81–86.

Kasari, C., Gulsrud, A. C., Wong, C., Kwon, S., & Locke, J. (2010). Randomized controlled caregiver mediated joint engagement intervention for toddlers with autism. *Journal of Autism and Developmental Disorders*, *40*(9), 1045–1056.

Kasari, C., Lawton, K., Shih, W., Barker, T. V., Landa, R., Lord, C., … & Senturk, D. (2014). Caregiver-mediated intervention for low-resourced preschoolers with autism: An RCT. *Pediatrics*, *134*(1), 72-79.

Kimhi, Y. (2014). Theory of mind abilities and deficits in autism spectrum disorders. *Topics in Language Disorders*, *34*(4), 329–343.

Kovács, Á. M., Téglás, E., & Endress, A. D. (2010). The social sense: Susceptibility to others' beliefs in human infants and adults. *Science*, *330*(6012), 1830–1834.

Leekam, S., Baron-Cohen, S., Perrett, D., Milders, M., & Brown, S. (1997). Eye-direction detection: A dissociation between geometric and joint attention skills in autism. *British Journal of Developmental Psychology*, *15*(1), 77–95.

Leekam, S. R., & Ramsden, C. A. H. (2006). Dyadic orienting and joint attention in preschool children with autism. *Journal of Autism and Developmental Disorders*, *36*(2), 185-197.

Liebal, K., Colombi, C., Rogers, S. J., Warneken, F., & Tomasello, M. (2008). Helping and cooperation in children with autism. *Journal of Autism and Developmental Disorders*, *38*(2), 224–238.

Lord, C., Rutter, M., DiLavore, P. C., Risi, S., Gotham, K., & Bishop, S. L. (2012). *Autism Diagnostic Observation Schedule-2*. Los Angeles: Western Psychological Services.

Luckett, T., Bundy, A., & Roberts, J. (2007). Do behavioural approaches teach children with autism to play or are they pretending? *Autism*, *11*(4), 365–388.

Meltzoff, A. N. (1995). Understanding the intentions of others: Re-enactment of intended acts by 18-month-old children. *Developmental Psychology*, *31*(5), 838–850.

Milligan, K., Astington, J. W., & Dack, L. A. (2007). Language and theory of mind: Meta-analysis of the relation between language ability and false-belief understanding. *Child Development*, *78*(2), 622–646.

Moss, J., & Howlin, P. (2009). Autism spectrum disorders in genetic syndromes: Implications for diagnosis, intervention and understanding the wider autism spectrum disorder population. *Journal of Intellectual Disability Research*, *53*(10), 852–873.

Mullen, E. M. (1995). *Mullen Scales of Early Learning.* Minneapolis: Pearson Assessments.

Nation, K., & Penny, S. (2008). Sensitivity to eye gaze in autism: Is it normal? Is it automatic? Is it social? *Development and Psychopathology*, *20*(1), 79–97.

NICE (2011). Autism spectrum disorder in under 19s: recognition, referral and diagnosis. Retrieved 6 April 2018, from https://www.nice.org.uk/guidance/cg128/chapter/Recommendations#referring-children-and-young-people-to-the-autism-team

Olineck, K. M., & Poulin-Dubois, D. (2005). Infants' ability to distinguish between intentional and accidental actions and its relation to internal state language. *Infancy*, *8*(1), 91–100.

Peterson, C. C., Wellman, H. M., & Liu, D. (2005). Steps in theory-of-mind development for children with deafness or autism. *Child Development*, *76*(2), 502–517.

Poslawsky, I. E., Naber, F. B., Bakermans-Kranenburg, M. J., van Daalen, E., van Engeland, H., & van IJzendoorn, M. H. (2015). Video-feedback Intervention to promote Positive Parenting adapted to Autism (VIPP-AUTI): A randomized

controlled trial. *Autism: The International Journal of Research and Practice*, *19*(5), 588–603.

Powis, L. (2014). Rubinstein-Taybi syndrome: From behaviour to cognition. Doctoral dissertation, University of Birmingham.

Premack, D., & Woodruff, G. (1978). Does the chimpanzee have a theory of mind? *Behavioral and Brain Sciences*, *1*(4), 515–526.

Scott, R. M., Baillargeon, R., Song, H., & Leslie, A. M. (2010). Attributing false beliefs about non-obvious properties at 18 months. *Cognitive Psychology*, *61*(4), 366–395.

Southgate, V., Senju, A., & Csibra, G. (2007). Action anticipation through attribution of false belief by 2-year-olds. *Psychological Science*, *18*(7), 587–592.

Surian, L., Caldi, S., & Sperber, D. (2007). Attribution of beliefs by 13-month-old infants. *Psychological Science*, *18*(7), 580–586.

Svetlova, M., Nichols, S. R., & Brownell, C. A. (2010). Toddlers' prosocial behavior: From instrumental to empathic to altruistic helping. *Child Development*, *81*(6), 1814–1827.

Tomasello, M., Carpenter, M., Call, J., Behne, T., & Moll, H. (2005). Understanding and sharing intentions: The origins of cultural cognition. *The Behavioral and Brain Sciences*, *28*(5), 675-691-735.

Warneken, F., Chen, F., & Tomasello, M. (2006). Cooperative activities in young children and chimpanzees. *Child Development*, *77*(3), 640–663.

Warneken, F., & Tomasello, M. (2006). Altruistic helping in human infants and young chimpanzees. *Science*, *311*(5765), 1301–1303.

Warneken, F., & Tomasello, M. (2007). Helping and cooperation at 14 months of age. *Infancy*, *11*(3), 271–294.

Wellman, H. M., Cross, D., & Watson, J. (2001). Meta-analysis of theory-of-mind development: The truth about false belief. *Child Development*, *72*(3), 655–684.

Wellman, H. M., Fuxi, F., & Peterson, C. C. (2011). Sequential progressions in a theory of mind scale: Longitudinal perspectives. *Child Development*, *82*(3), 780–792.

Wellman, H. M., & Liu, D. (2004). Scaling of theory-of-mind tasks. *Child Development*, *75*(2), 523–541.

Wellman, H. M., Lopez-Duran, S., LaBounty, J., & Hamilton, B. (2008). Infant attention to intentional action predicts preschool theory of mind. *Developmental Psychology*, *44*(2), 618–623.

CHAPTER 3

Public dissemination document

1. Introduction

Autism spectrum disorder (ASD) is a neurodevelopmental disorder estimated to occur in approximately 1% of the general population (Baird et al., 2006; Baron-Cohen et al., 2009). ASD is characterised by communication difficulties, impairments in social interaction, and the presence of restricted and repetitive interests and behaviours (American Psychiatric Association, 2013; World Health Organisation, 1992). The following describes two studies that aim to further our understanding of genetic, cognitive and developmental processes in ASD.

2. Chapter 1: Exploring sources of variance in autism prevalence in rare syndromes: A meta-analysis

2.1 Background

The genetic causes of ASD are still unknown, and evidence suggests that ASD likely arises from complex interactions between multiple genetic and environmental factors (Persico & Bourgeron, 2006; Zhao et al., 2007). One approach to teasing apart these risk factors is to study genetic syndromes in which ASD is more common; because we know more about the genetic causes of these syndromes, studying the links between genes and behaviour in these groups might help us to understand the genetics of ASD more broadly.

To help focus these research efforts, Richards, Jones, Groves, Moss and Oliver (2015) carried out a research review and generated prevalence estimates for 12 syndromes associated with higher rates of ASD. This research concluded ASD was highly prevalent in tuberous sclerosis complex and Rett, Cohen, Cornelia de Lange, Angelman and CHARGE

syndromes. Although these findings provided some direction for research into gene-behaviour links, Richards et al. (2015) reported that ASD prevalence estimates varied greatly between different studies. This makes it more difficult to produce reliable overall estimates of ASD prevalence in each syndrome, and suggests that participant characteristics *other than* syndrome diagnosis influenced reported rates of ASD. Evidence suggests that the severity and prevalence of ASD in the general population varies according to age and gender (Baird et al., 2006; Woodman, Smith, Greenberg, & Mailick, 2015), therefore it is possible that these factors would contribute to differences in ASD prevalence estimates in different studies. Furthermore, lower intellectual ability on its own is associated with higher rates of ASD; given that intellectual disability is a common feature of many genetic syndromes, it is possible that higher rates of ASD in genetic syndromes are a result of lower intellectual ability, rather than syndrome-specific factors per se (Moss & Howlin, 2009). Finally, the ways in which studies were carried out might also influence ASD prevalence estimates. Studies used different tools for assessing ASD, for example, which have different levels of accuracy for detecting ASD traits. The extent to which these differences influenced ASD prevalence estimates was not addressed by Richards et al. (2015).

On the basis of this prior research the current study had two aims. The first was to review the most recent research in order to provide up to date estimates of ASD prevalence in genetic syndromes. The second was to explore the extent to which ASD prevalence estimates were influenced by study factors (ASD assessment method, the method used to confirm genetic syndrome diagnoses, and recruitment strategies) and participant characteristics (intellectual ability, age and gender).

2.2 Method

Computerised research databases were searched for papers reporting ASD prevalence estimates in 21 genetic syndromes associated with elevated rates of ASD. Studies were rated against quality criteria produced by Richards et al. (2015), and studies with low quality-ratings were excluded from the analysis. Statistical analyses produced pooled estimates of ASD prevalence in each genetic syndrome and evaluated the impact of study factors and participant characteristics on ASD prevalence.

2.3 Results

The literature search identified 14 syndromes with sufficient, good-quality research reporting ASD prevalence estimates. Results showed that the chances of an individual having ASD were higher for all 14 syndromes compared to the general population. High prevalence estimates (greater than 30%) in tuberous sclerosis complex and Rett, Cohen, Angelman, Sotos and Cornelia de Lange syndromes confirmed that these syndromes would be a useful focus for research exploring gene-behaviour links in ASD.

ASD prevalence estimates were significantly affected by ASD assessment method in nine syndromes, by recruitment strategy in four syndromes, and by the method through which syndrome diagnoses were confirmed in six syndromes. These results showed that study factors had a significant impact on ASD prevalence estimates, demonstrating the importance of carrying out good-quality studies with appropriate ASD assessment tools.

When the data for all syndromes were combined, higher rates of ASD were found in samples reporting a higher proportion of individuals with an intellectual disability and

lower average IQ. Thus whilst a number of syndrome-specific diagnoses result in higher rates of ASD, this might be attributed, to some extent, to lower intellectual ability in these groups. Participant age had a mixed effect on ASD rates, suggesting the relationship between ASD severity and age may be complex. There was no relationship between gender and ASD prevalence, indicating that gender has minimal influence on rates of ASD in these groups.

2.4 Clinical implications

In clinical settings there is a risk that additional diagnoses are overlooked in individuals with an existing genetic syndrome diagnosis. The findings from this study, which evidence higher rates of ASD in all syndromes assessed, highlight the importance of ASD assessment where appropriate, to ensure appropriate clinical and educational intervention. However, the evidence also indicates that intellectual ability, age and ASD assessment tools influence ASD classification in genetic syndromes. ASD diagnoses should thus be made on the basis of careful assessment and formulation, carried out by individuals with specialist knowledge of the behavioural profiles typical of these groups.

3. Chapter 2: A developmental scale of early social cognition in autism spectrum disorder

3.1 Background

'Theory of mind' (ToM), a term coined by Premack and Woodruff (1978), refers to the ability to understand the mental states of others, and allows individuals to engage in complex social behaviours (Baron-Cohen, Tager-Flusberg, & Lombardo, 2013). Evidence

has consistently shown that children have difficulty understanding *false belief* (i.e. an understanding that another person can hold a belief that the individual knows to be untrue) until 4-5 years of age (Wellman, Cross, & Watson, 2001). This has been interpreted as evidence that younger children 'lack' ToM. More recent research, however, suggests that ToM abilities are present at a much younger age. Powis (2014) showed that typically developing children aged 14-34 months understood other people's *intentions,* and engaged in *shared intentionality*, referring to the ability to engage jointly with another in pursuit of a shared goal. Further 'scaling' analyses, which look at whether there are patterns in the order in which success on different tasks is achieved, suggested that children acquired these skills in a consistent sequence. This indicated that earlier skills might act as precursors to later-developing abilities.

Evidence suggests that children with ASD are delayed in their development of more complex ToM abilities, such as understanding *belief* and *desire,* and that these deficits might underpin their difficulties in social interaction and communication (Baron-Cohen, 1995). However, the development of early ToM skills, such as understanding *intention* and *shared intentionality* are less understood in ASD, and this formed the basis for the current research.

3.2 Method

Twenty-one children with ASD were recruited from a school and university research centre in the West Midlands. Children were assessed using Powis' (2014) Early Social Cognition Scale, a collection of six tasks assessing understanding of others' *intentions* and the ability to engage in *shared intentionality*.

3.3 Results

Scaling analyses revealed that children with ASD developed early ToM abilities in a consistent sequence, which paralleled that observed in Powis' (2014) sample of typically developing children. The consistency between these sequences is a further indication that earlier-developing skills might act as a foundation for later-developing abilities.

Despite these parallels in developmental *sequence*, the data suggested that children with ASD were delayed in the development of early ToM abilities when compared with typically developing children. Furthermore, interpreting others' intentions through eye gaze was a particularly challenging task for children with ASD. Children with ASD might have difficulty *understanding* eye gaze as a communicative cue. Alternatively, children with ASD might miss these communicative cues due to a failure to *attend* to others' eyes (Nation & Penny, 2008). Finally, although some children could coordinate their actions with another person in a problem-solving task, no children with ASD engaged in a similar task requiring them to engage jointly with another as part of a social game. Evidence suggests that children with ASD are less *motivated* to initiate and maintain social interaction (Chevallier, Kohls, Troiani, Brodkin, & Schultz, 2012), which might explain their difficulties in performing this task.

3.4 Clinical implications

Understanding the profile of social cognitive skills in ASD has implications for assessment, helping to differentiate the processes underlying behaviours seen in ASD and other neurodevelopmental disorders to ensure appropriately targeted intervention. Findings suggested that interpreting intention in eye gaze and engaging in cooperative games were

both challenging skills for children with ASD, and that these skills might be necessary for later ToM development. As such, approaches that support the development of these skills could be important targets for clinical intervention.

4. References

American Psychiatric Association (2013). *Diagnostic and Statistical Manual of Mental Disorders.* (5th ed.). Arlington, VA: American Psychiatric Publishing.

Baird, G., Simonoff, E., Pickles, A., Chandler, S., Loucas, T., Meldrum, D., & Charman, T. (2006). Prevalence of disorders of the autism spectrum in a population cohort of children in South Thames: the Special Needs and Autism Project (SNAP). *The Lancet*, *368*(9531), 210–215.

Baron-Cohen, S. (1995). *Mindblindness: An Essay on Autism and Theory of Mind.* Oxford: MIT Press.

Baron-Cohen, S., Scott, F. J., Allison, C., Williams, J., Bolton, P., Matthews, F. E., & Brayne, C. (2009). Prevalence of autism-spectrum conditions: UK school-based population study. *The British Journal of Psychiatry: The Journal of Mental Science*, *194*(6), 500–509.

Baron-Cohen, S., Tager-Flusberg, H., & Lombardo, M. (Eds.). (2013). *Understanding Other Minds: Perspectives from Developmental Social Neuroscience*. New York: Oxford University Press.

Chevallier, C., Kohls, G., Troiani, V., Brodkin, E. S., & Schultz, R. T. (2012). The social motivation theory of autism. *Trends in Cognitive Sciences*, *16*(4), 231–239.

Moss, J., & Howlin, P. (2009). Autism spectrum disorders in genetic syndromes: Implications for diagnosis, intervention and understanding the wider autism spectrum disorder population. *Journal of Intellectual Disability Research*, *53*(10), 852–873.

Nation, K., & Penny, S. (2008). Sensitivity to eye gaze in autism: Is it normal? Is it automatic? Is it social? *Development and Psychopathology*, *20*(1), 79–97.

Persico, A. M., & Bourgeron, T. (2006). Searching for ways out of the autism maze: Genetic, epigenetic and environmental clues. *Trends in Neurosciences*, *29*(7), 349–358.

Powis, L. (2014). Rubinstein-Taybi syndrome: From behaviour to cognition. Doctoral dissertation, University of Birmingham.

Premack, D., & Woodruff, G. (1978). Does the chimpanzee have a theory of mind? *Behavioral and Brain Sciences*, *1*(4), 515–526.

Richards, C., Jones, C., Groves, L., Moss, J., & Oliver, C. (2015). Prevalence of autism spectrum disorder phenomenology in genetic disorders: A systematic review and meta-analysis. *The Lancet Psychiatry*, *2*(10), 909–916.

Wellman, H. M., Cross, D., & Watson, J. (2001). Meta-analysis of theory-of-mind development: The truth about false belief. *Child Development*, *72*(3), 655–684.

Woodman, A. C., Smith, L. E., Greenberg, J. S., & Mailick, M. R. (2015). Change in autism symptoms and maladaptive behaviors in adolescence and adulthood: The role of positive family processes. *Journal of Autism and Developmental Disorders*, *45*(1), 111–126.

World Health Organisation (1992). *ICD-10 Classifications of Mental and Behavioural Disorder: Clinical Descriptions and Diagnostic Guidelines*. Geneva: World Health Organisation.

Zhao, X., Anthony, L., Vlad, K., Clara, L., Daniel, G., Kiely, L., …& Michael, W. (2007). A unified genetic theory for sporadic and inherited autism. *Proceedings of the National Academy of Sciences*, *104*(31), 12831–12836.

Appendix 1: Syndrome groups, search dates and search terms included in the literature search

	PsychINFO		MEDLINE		Embase		PubMed Central		Search terms
	Date searched	**Inclusion dates**	**Date searched**	**Inclusion dates**	**Date searched**	**Inclusion dates**	**Date searched**	**Inclusion dates**	
Fragile X syndrome (FraX)	17/08/17	2014 to August Week 2 2017	17/08/17	2014 to August Week 2 2017	17/08/17	2014 – 17th August 2017	14/10/17	01/02/14 to 14/10/17	Fragile X; Fragile-X; Fragile X syndrome; FXS; FRAXA syndrome; AFRAX; Martin-Bell* syndrome; Marker X syndrome; fraX syndrome; fra(X) syndrome; X-linked mental retardation; Macroorchidism; Escalante* syndrome; Escalante*
Tuberous Sclerosis Complex (TSC)	25/08/17	2014 to August Week 3 2017	25/08/17	2014 to August Week 2 2017	25/08/17	2014 to 24th August 2017	01/09/17	01/02/14 to 01/09/17	Tuberous sclerosis; Tuberous sclerosis syndrome; Bourneville* disease; Bourneville* phakomatosis; Cerebral sclerosis; Cerebral sclerosis syndrome; Epiloia; Sclerosis tuberose; Tuberose sclerosis; Tuberose sclerosis syndrome; Tuberous sclerosis complex; TSC; TSS
Rett's syndrome (Rett)	25/08/17	2014 to August Week 3 2017	25/08/17	2014 to August Week 2 2017	25/08/17	2014 to 24th August 2017	04/10/17	01/02/14 to 04/10/17	Rett*; Rett* syndrome; Rett* disorder; RTS; RTT; Cerebroatrophic hyperammonemia; Autism-dementia-ataxia-loss of purposeful hand use syndrome
Down syndrome (DS)	17/08/17	2014 to August Week 2 2017	17/08/17	2014 to August Week 2 2017	17/08/17	2014 to 17th August 2017	14/10/17	01/02/17 to 14/10/17	Down*; Down* syndrome; Trisomy 21; Trisomy G; 47,XX,+21; 47,XY,+2

	PsychINFO		MEDLINE		Embase		PubMed Central		Search terms
	Date searched	**Inclusion dates**	**Date searched**	**Inclusion dates**	**Date searched**	**Inclusion dates**	**Date searched**	**Inclusion dates**	
Phenylketonuria syndrome (PKU)	25/08/17	2014 to August Week 3 2017	25/08/17	2014 to August Week 2 2017	25/08/17	2014 to 24/08/17	03/10/17	01/02/14 to 03/10/17	Phenylketonuria; Phenylalanine hydroxylase; Folling* disease; Folling* syndrome; PAH deficiency; PAH deficiency disease; Phenylalanine hydroxylase deficiency disease; Phenylalanine hydroxylase deficiency; PKU; Oligophrenia phenylpyruvica; Deficiency Disease, Phenylalanine Hydroxylase
CHARGE syndrome (CHARGE)	17/08/17	2014 to August Week 2 2017	17/08/17	2014 to August Week 2 2017	17/08/17	2014 to 17/08/17	04/10/17	01/02/17 to 04/10/17	CHARGE; CHARGE syndrome; CHARGE association; Hall-Hittner* syndrome; Hall* Hittner* syndrome
Angelman syndrome (AS)	17/08/17	2014 to August Week 2 2017	17/08/17	2014 to August Week 2 2017	17/08/17	2014 to 17/08/17	01/09/17	01/02/14 to 01/09/17	Angelman*; Angelman* syndrome; AS; Happy puppet syndrome; Happy puppet
Neurofibromatosis Type 1 (NF1)[1]	17/08/17	2014 to August Week 2 2017	17/08/17	2014 to August Week 2 2017	17/08/17	2014 to 17th August 2017	04/10/17	01/02/14 to 04/10/17	Neurofibromatosis; Neurofibromatosis type 1; Neurofibromatosis 1; NF1; Peripheral Neurofibromatosis; Recklinghausen* disease; Neurofibromatosis type 2; Neurofibromatosis 2; NF2; Central neurofibromatosis; Bilateral acoustic neurofibromatosis; BANF; Familial acoustic neuromas

[1] The search included terms for both Neurofibromatosis type 1 and type 2, however, no papers reporting ASD prevalence estimates in Neurofibromatosis type 2 met specified inclusion criteria

	PsychINFO		MEDLINE		Embase		PubMed Central		Search terms
	Date searched	**Inclusion dates**	**Date searched**	**Inclusion dates**	**Date searched**	**Inclusion dates**	**Date searched**	**Inclusion dates**	
Joubert syndrome (JS)	17/08/17	2014 to August Week 2 2017	17/08/17	2014 to August Week 2 2017	17/08/17	2014 to 17/08/17	02/10/17	01/02/17 to 02/10/17	Joubert*; Joubert* syndrome; Joubert-Bolthauser* syndrome; JBTS; Cerebello-oculo-renal syndrome; Cerebello-oculo-renal syndrome 1; Cerebellooculorenal syndrome 1; Cerebellooculorenal syndrome; CORS; CORS1; Cerebellar vermis agenesis; Cerebelloparenchymal disorder 4; Cerebelloparenchymal disorder; CPD; CPD4; Familial aplasia of the vermis
William's syndrome (WS)	25/08/17	2014 to August Week 3 2017	25/08/17	2014 to August Week 2 2017	25/08/17	2014 to 24/08/17	04/10/17	01/02/14 to 04/10/17	William*; William* syndrome; Beuren* syndrome; Elfin Facies syndrome; Hypercalcemia-Supravalvar Aortic Stenosis; Infantile hypercalcemia; Supravalvar aortic stenosis syndrome; WBS; Williams-Beuren* syndrome; WMS; WS
Goldenhar syndrome (GS)	17/08/17	2014 to August Week 2 2017	17/08/17	2014 to August Week 2 2017	17/08/17	2014 to 17/08/17	04/10/17	01/02/14 to 04/10/17	Goldenhar*; Goldenhar* syndrome; Oculoauriculovertebral spectrum; Oculoauriculovertebral syndrome; Oculoauriculovertebral dysplasia; OAV; OAVD; OAVS; Oculo-Auriculo-Vertebral syndrome; Oculo-Auriculo-Vertebral spectrum; Oculo-Auriculo-Vertebral dysplasia; Brachial arch syndrome; Facioauriculovertebral syndrome; FAV; FAVS; Lateral facial dysplasia

	PsychINFO		MEDLINE		Embase		PubMed Central		Search terms
	Date searched	**Inclusion dates**	**Date searched**	**Inclusion dates**	**Date searched**	**Inclusion dates**	**Date searched**	**Inclusion dates**	
Hypomelanosis of Ito syndrome (HoI)	17/08/17	2014 to August Week 2 2017	17/08/17	2014 to August Week 2 2017	17/08/17	2014 to 17/08/17	04/10/17	01/02/14 to 04/10/17	Hypomelanosis of Ito; Ito hypomelanosis; Incontinentia pigmentosa achromians; Ito syndrome; ITO; IPA; HMI
Noonan syndrome (Noonan)	17/08/17	2014 to August Week 2 2017	17/08/17	2014 to August Week 2 2017	17/08/17	2014 to 17/08/17	04/10/17	01/02/14 to 04/10/17	Noonan*; Noonan* syndrome; Nunan*; Nunan* syndrome; Familial Turner* syndrome; Female pseudo-Turner syndrome; Male Turner* syndrome; Noonan-Ehmke* syndrome; Nunan-Ehmke* syndrome; Pseudo-Ullrich-Turner* syndrome; Turner-like syndrome; Ullrich-Noonan* syndrome; Ullrich-Nunan* syndrome; Turner* phenotype, karyotype normal; Turner syndrome in female with X chromosome
Sotos syndrome (Sotos)	25/08/17	2014 to August Week 3 2017	25/08/17	2014 to August Week 2 2017	25/08/17	2014 to 24/08/17	04/10/17	01/02/17 to 04/10/17	Sotos*; Sotos* syndrome; Cerebral gigantism; Sotos* sequence
Leber's Amaurosis syndrome (Leber's)	17/08/17	2014 to August Week 2 2017	17/08/17	2014 to August Week 2 2017	17/08/17	2014 to 17/08/17	04/10/17	01/02/14 to 04/10/17	Leber* amaurosis; Leber* congenital amaurosis; LCA; Congenital retinal blindness; CRB; Dysgenesis neuroepithelialis retinae; Hereditary epithelial dysplasia of retina; Hereditary retinal aplasia; Heredoretinopathia congenitalis; Leber* abiotrophy; Leber* congenital tapetoretinal degeneration

	PsychINFO		MEDLINE		Embase		PubMed Central		Search terms
	Date searched	**Inclusion dates**	**Date searched**	**Inclusion dates**	**Date searched**	**Inclusion dates**	**Date searched**	**Inclusion dates**	
22q11.2 deletion syndrome (22q11.2)	22/09/17	2014 to September Week 3 2017	22/09/17	2014 to September Week 2 2017	22/09/17	2014 to 17/08/17	04/10/17	01/02/14 to 04/10/17	VCF; VCFS; Velocardiofacial syndrome; CTAF; Velo-cardio-facial syndrome; DiGeorge* syndrome; Conotruncal anomaly face syndrome; CATCH22; Autosomal dominant Opitz G/BBB syndrome; Autosomal dominant Opitz G BBB syndrome; Cayler cardiofacial syndrome; Deletion 22q11/2 syndrome; 22q11/2 deletion syndrome; 22q11/2DS; 22q11 deletion syndrome; Sedlackova* syndrome; Shprintzen* syndrome
Cohen syndrome (Cohen)	17/08/17	2014 to August Week 2 2017	17/08/17	2014 to August Week 2 2017	17/08/17	2014 to 17/08/17	04/10/17	01/02/14 to 04/10/17	Cohen* syndrome; Norio* syndrome; Obesity-hypotonia syndrome; Pepper* syndrome; Prominent incisors-obesity-hypotonia syndrome; Hypotonia obesity and prominent incisors
Cornelia de Lange syndrome (CdLS)	22/09/17	2014 to September Week 3 2017	22/09/17	2014 to September Week 2 2017	22/09/17	2014 to 21/09/17	04/10/17	01/02/14 to 04/10/17	Cornelia de Lange* syndrome; CDLS; De Lange* syndrome; Branchmann-De Lange* syndrome; BDLS; Brachmann* syndrome; Amstelodamensis typus degenerativus; Amsterdam dwarf syndrome; Amsterdam dwarfism

	PsychINFO		MEDLINE		Embase		PubMed Central		Search terms
	Date searched	**Inclusion dates**	**Date searched**	**Inclusion dates**	**Date searched**	**Inclusion dates**	**Date searched**	**Inclusion dates**	
Ehlers-Danlos syndrome (EDS)	17/08/17	2014 to August Week 2 2017	17/08/17	2014 to August Week 2 2017	17/08/17	2014 to 17/08/17	04/10/17	01/02/14 to 04/10/17	Ehlers-Danlos; Ehlers-Danlo*; Ehlers-Danlos syndrome; Ehlers-Danlo* syndrome; EDS; Ehlers-Danlos disease; Ehlers-Danlo* disease; Ehlers Danlos; Ehler* Danlo*; Ehlers Danlos syndrome; Ehler* Danlo* syndrome; Ehlers Danlos disease; Ehler* Danlo* disease; ED syndrome; vascular-Ehler* Danlo* syndrome; vascular ehler* danlo* syndrome; vascular ehler* danlo*; vascular-Ehler* Danlo*; vEDS
Lujan-Fryns syndrome (LFS)	17/08/17	2014 to August Week 2 2017	17/08/17	2014 to August Week 2 2017	17/08/17	2014 to 17/08/17	04/10/17	01/02/14 to 04/10/17	Lujan-Fryns*; Lujan-Fryn*; Lujan-Fryns* syndrome; LFS; Lujan* syndrome; X-linked intellectual deficit with marfanoid habitus; X-linked intellectual deficit with marfanoid features; X-linked mental retardation with marfanoid features; X-linked mental retardation with marfanoid habitus; XLMR with marfanoid features; XLMR with marfanoid habitus
Moebius syndrome (Moebius)	17/08/17	1967 to February Week 3 2014	17/08/17	1946 to February Week 3 2014	17/08/17	1974 to 2014 Week 09	04/10/17	01/02/14 to 04/10/17	Moebius*; Mobius*; Moebius* syndrome; Mobius* syndrome; Moebius* spectrum; Mobius* spectrum; Moebius* sequence; Mobius* sequence; Congenital facial diplegia; Congenital ophthalmoplegia and facial paresis; Moebius* congenital oculofacial paralysis; Mobius* congenital oculofacial paralysis

Appendix 2: Papers included and excluded during selection for each syndrome

	Identification		Screening		Eligibility		Quality		Included
	Records identified through database searching	**Records after duplicates removed**	**Number of papers screened**	**Excluded**	**Full text papers assessed for eligibility**	**Excluded with reasons**	**Papers assessed for quality**	**Excluded**	**Papers included in meta-analysis**
FraX	4263	3342	3342	3301	41	28 [b,c,d,e,f]	13	4	9
22q11.2	1071	980	980	965	15	6 [b,c,d,g]	9	1	8
TSC	2175	1977	1977	1963	14	5 [d,e,h]	9	2	7
DS	8038	7160	7160	7146	14	10 [c,d,e]	4	0	4
NF1	519	415	415	405	10	6 [b,c,d,i,j]	4	2	2
PKU	274	229	229	225	4	2 [d]	2	0	2
Sotos	90	76	76	74	2	0	2	0	2
CdLS	276	245	245	241	4	3 [c,k]	1	0	1
AS	1081	943	943	940	3	2[d]	1	0	1
Noonan	616	600	600	598	2	1 [c]	1	0	1
WS	2392	2264	2264	2259	5	4 [c,d,e]	1	1	0
CHARGE	181	118	118	117	1	0	1	1	0
Rett	2471	2068	2068	2065	3	3 [c,h,i]	0	---	---
Cohen	1519	1506	1506	1505	1	1 [c]	0	---	---
E-D	137	103	103	102	1[a]	---	---	---	---
Moebius	12	11	11	11	0	---	---	---	---
JS	190	182	182	182	0	---	---	---	---
HoI	518	510	510	510	0	---	---	---	---
GS	184	183	183	183	0	---	---	---	---
L-F	67	64	64	64	0	---	---	---	---
Leber's	227	221	221	221	0	---	---	---	---
Total	**26301**	**23197**	**23197**	**23077**	**119**	**71**	**48**	**11**	**37**

[a] Fewer than two studies available for syndrome group (in combination with studies generated by Richards et al., 2015)
[b] Study reported on the same sample as another paper
[c] Study did not report the number/proportion of participants meeting clinical cut-off for ASD
[d] Participants recruited/excluded due to additional features, e.g. seizures, sensory impairments, self-injury, verbal ability
[e] Full text not available in English
[f] Study included participants with Fragile X pre-mutation
[g] Sample included 22q11.2 duplication syndrome
[h] Unable to access full text paper
[i] Participants recruited/excluded due to a previous or suspected ASD diagnosis
[j] Sample includes participants without full diagnosis
[k] Study included in Richards et al. (2015)

Appendix 3: Forest plots displaying random-effects and quality-effects pooled prevalence estimates for each syndrome (presented in descending order according to quality-weighted pooled prevalence estimate)

Rett syndrome

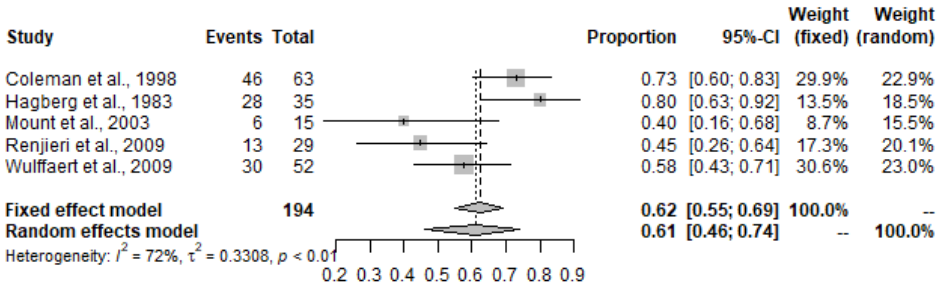

Figure 1. Pooled ASD prevalence estimates in Rett syndrome (random-effects model)

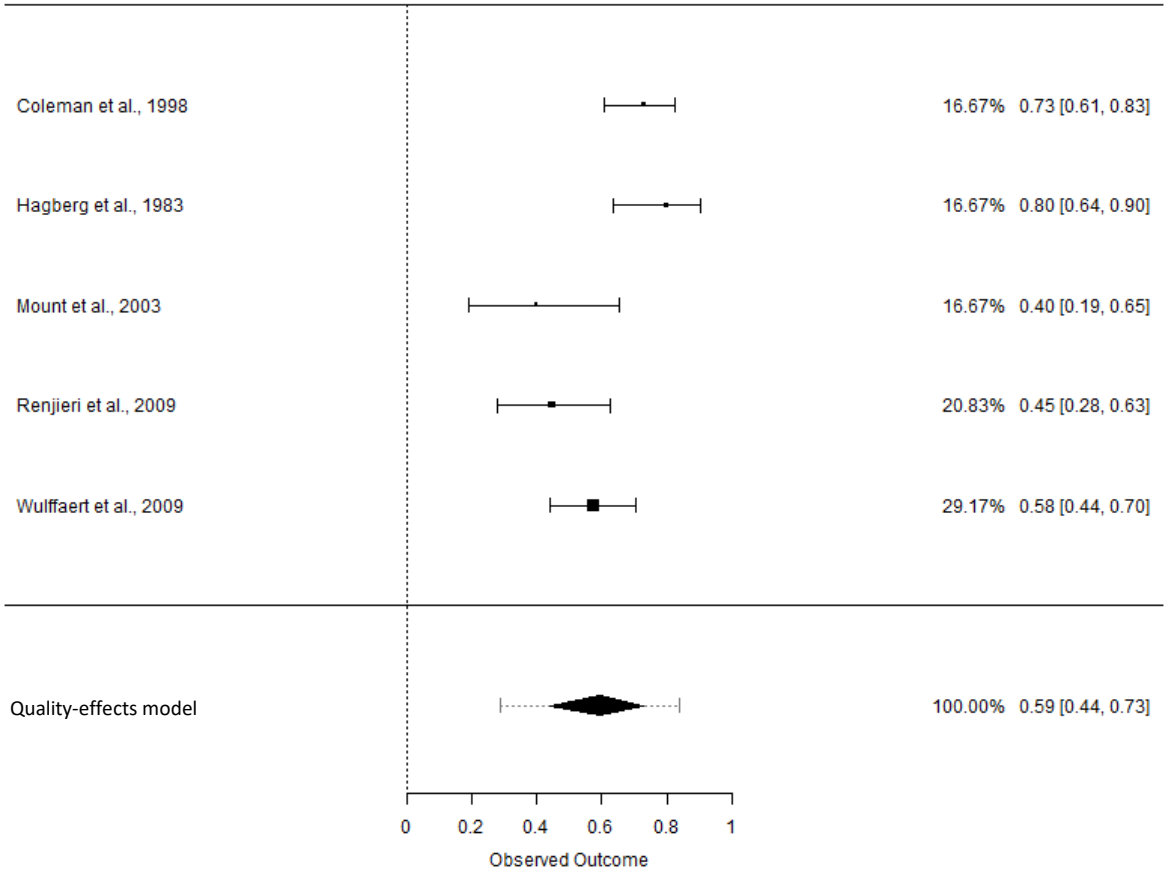

Figure 2. Pooled ASD prevalence estimates in Rett syndrome (quality-effects model)

Cohen syndrome

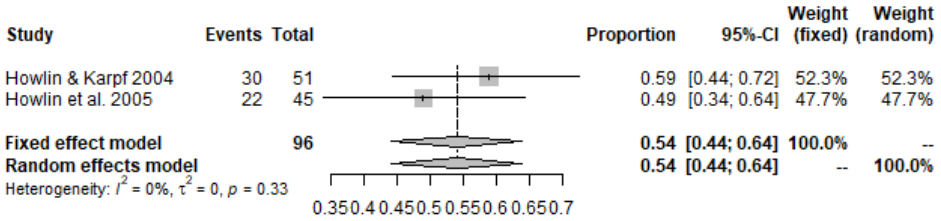

Figure 3. Pooled ASD prevalence estimates in Cohen syndrome (random-effects model)

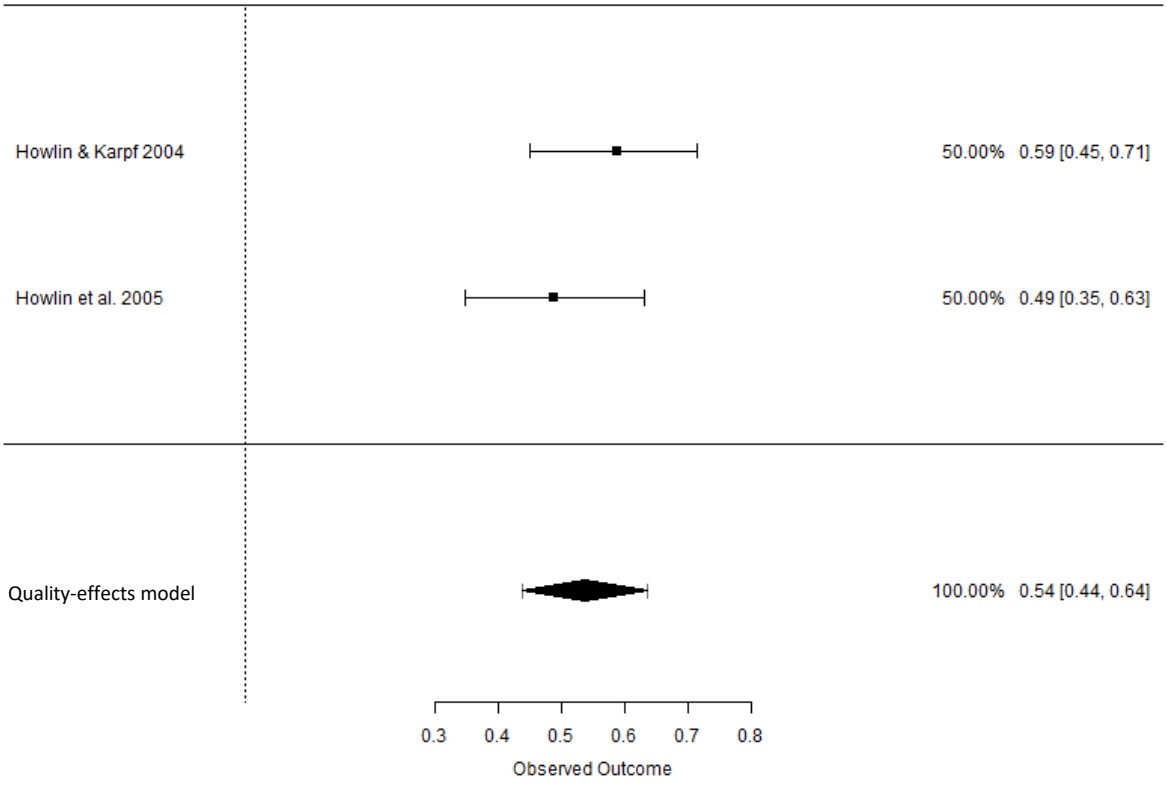

Figure 4. Pooled ASD prevalence estimates in Cohen syndrome (quality-effects model)

Angelman syndrome

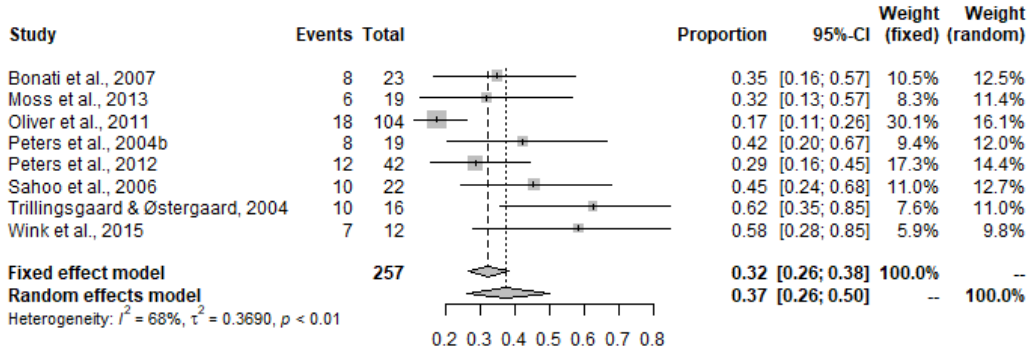

Figure 5. Pooled ASD prevalence estimates in Angelman syndrome (random-effects model)

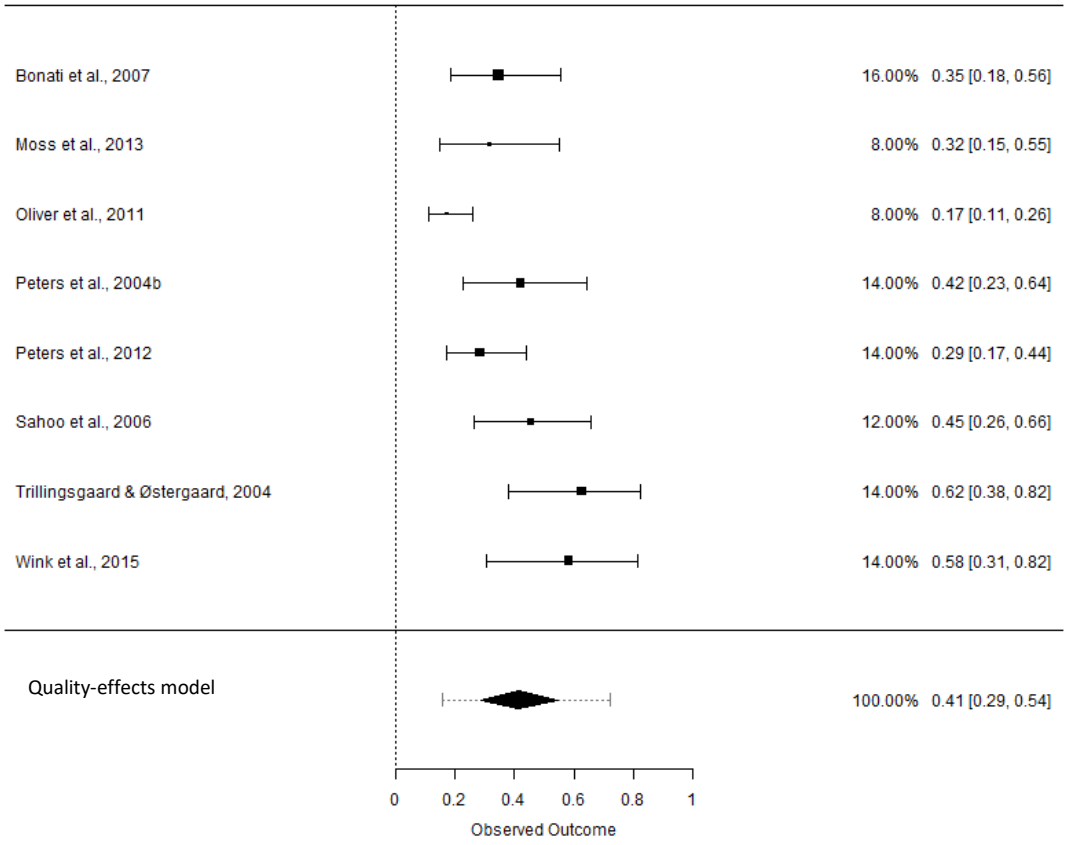

Figure 6. Pooled ASD prevalence estimates in Angelman syndrome (quality-effects model)

Sotos syndrome

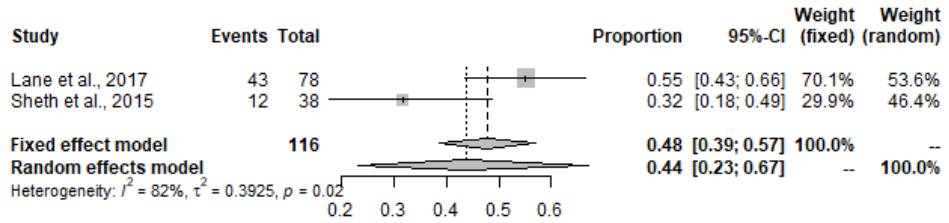

Figure 7. Pooled ASD prevalence estimates in Sotos syndrome (random-effects model)

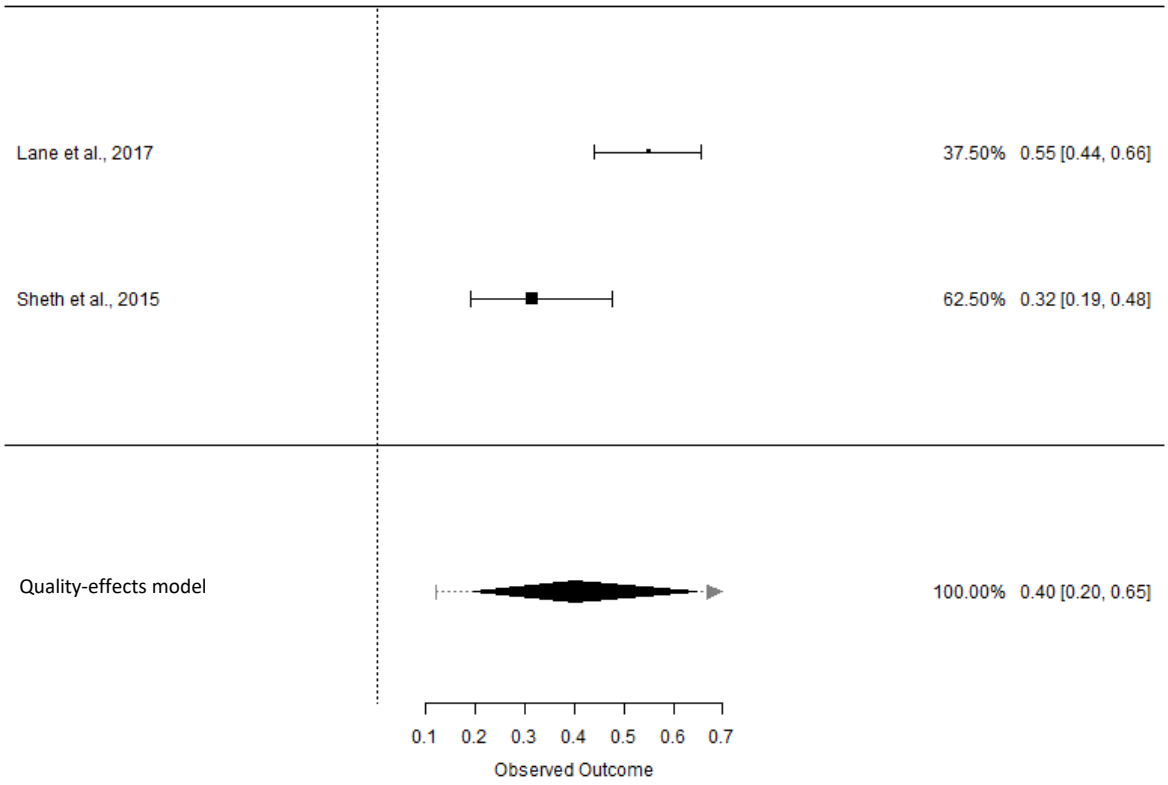

Figure 8. Pooled ASD prevalence estimates in Sotos syndrome (quality-effects model)

Cornelia de Lange syndrome

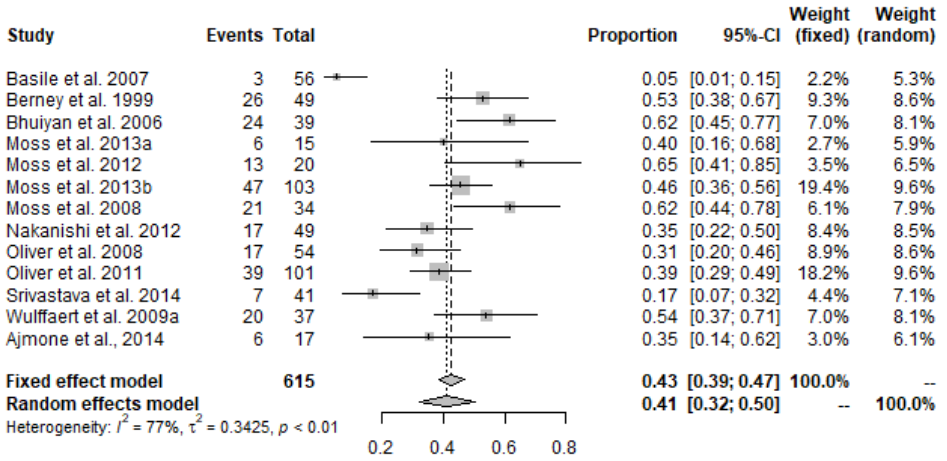

Figure 9. Pooled ASD prevalence estimates in Cornelia de Lange syndrome (random-effects model)

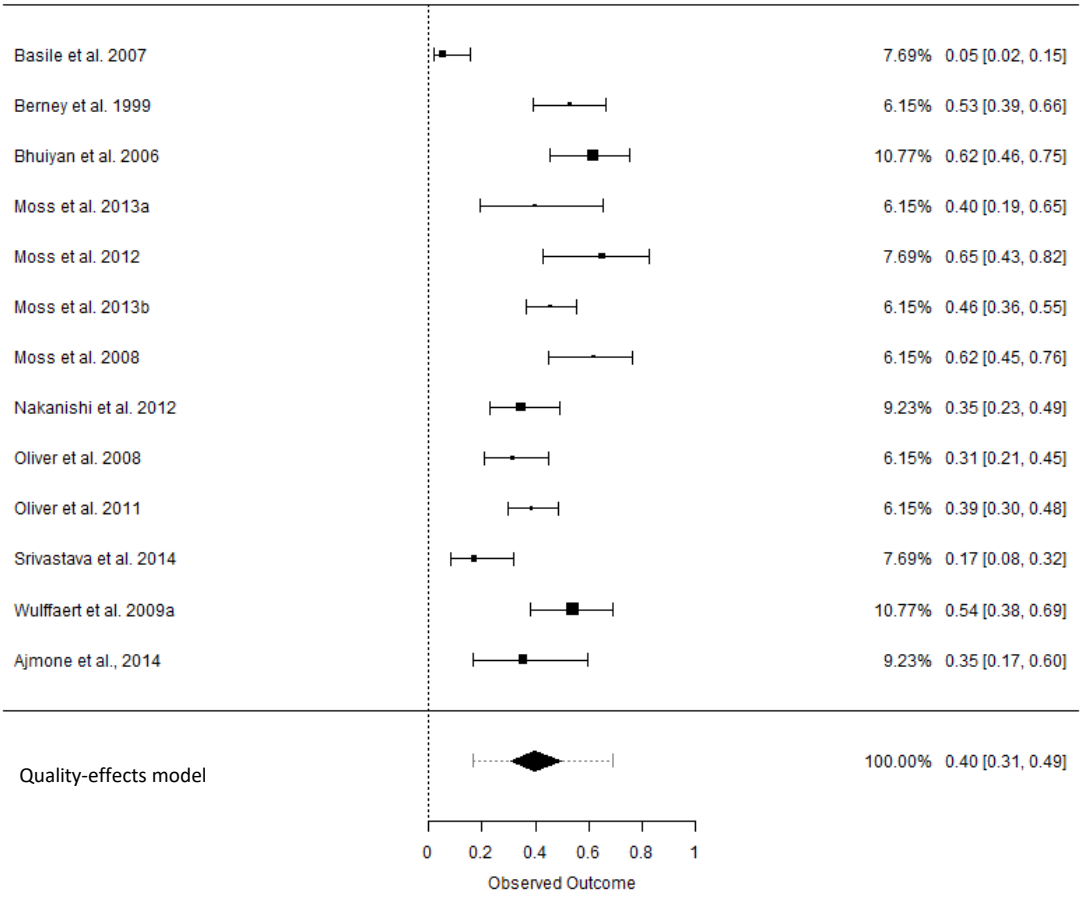

Figure 10. Pooled ASD prevalence estimates in Cornelia de Lange syndrome (quality-effects model)

Tuberous sclerosis complex

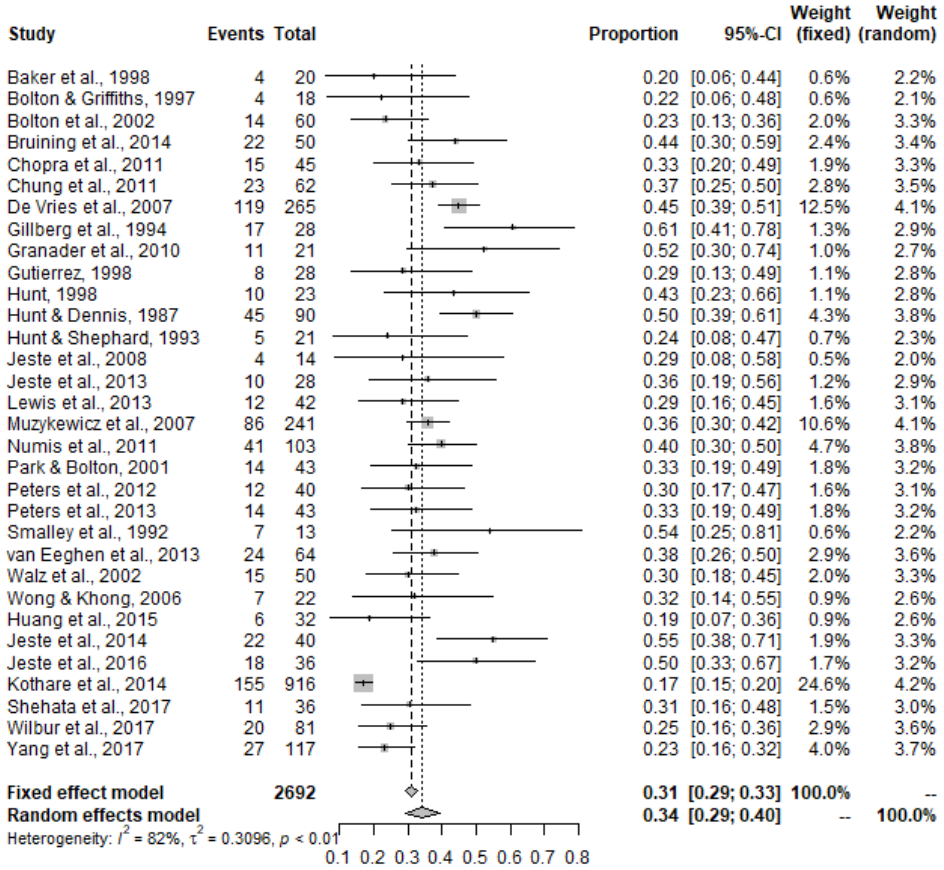

Figure 11. Pooled ASD prevalence estimates in tuberous sclerosis complex (random-effects model)

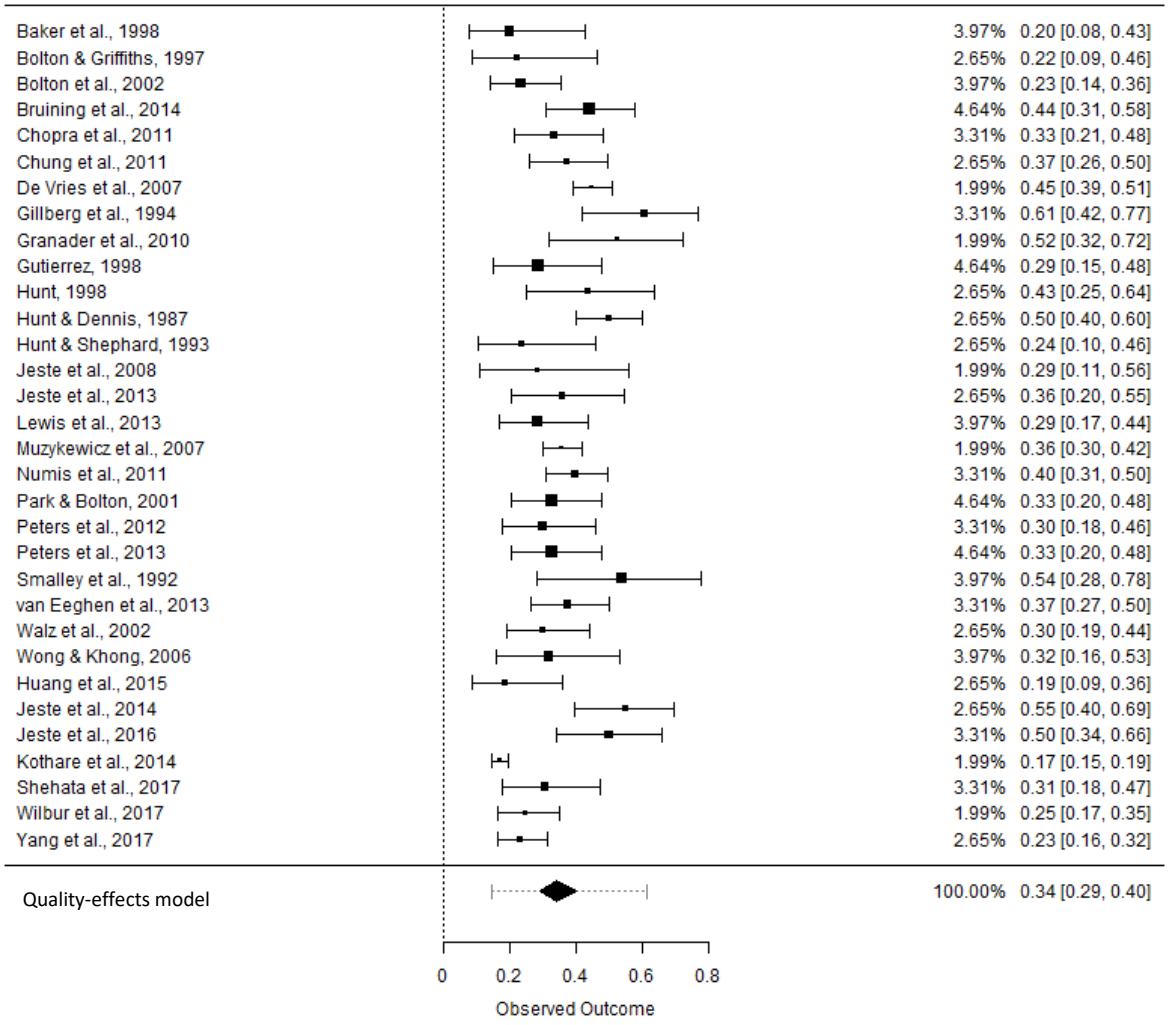

Figure 12. Pooled ASD prevalence estimates in tuberous sclerosis complex (quality-effects)

Fragile X syndrome

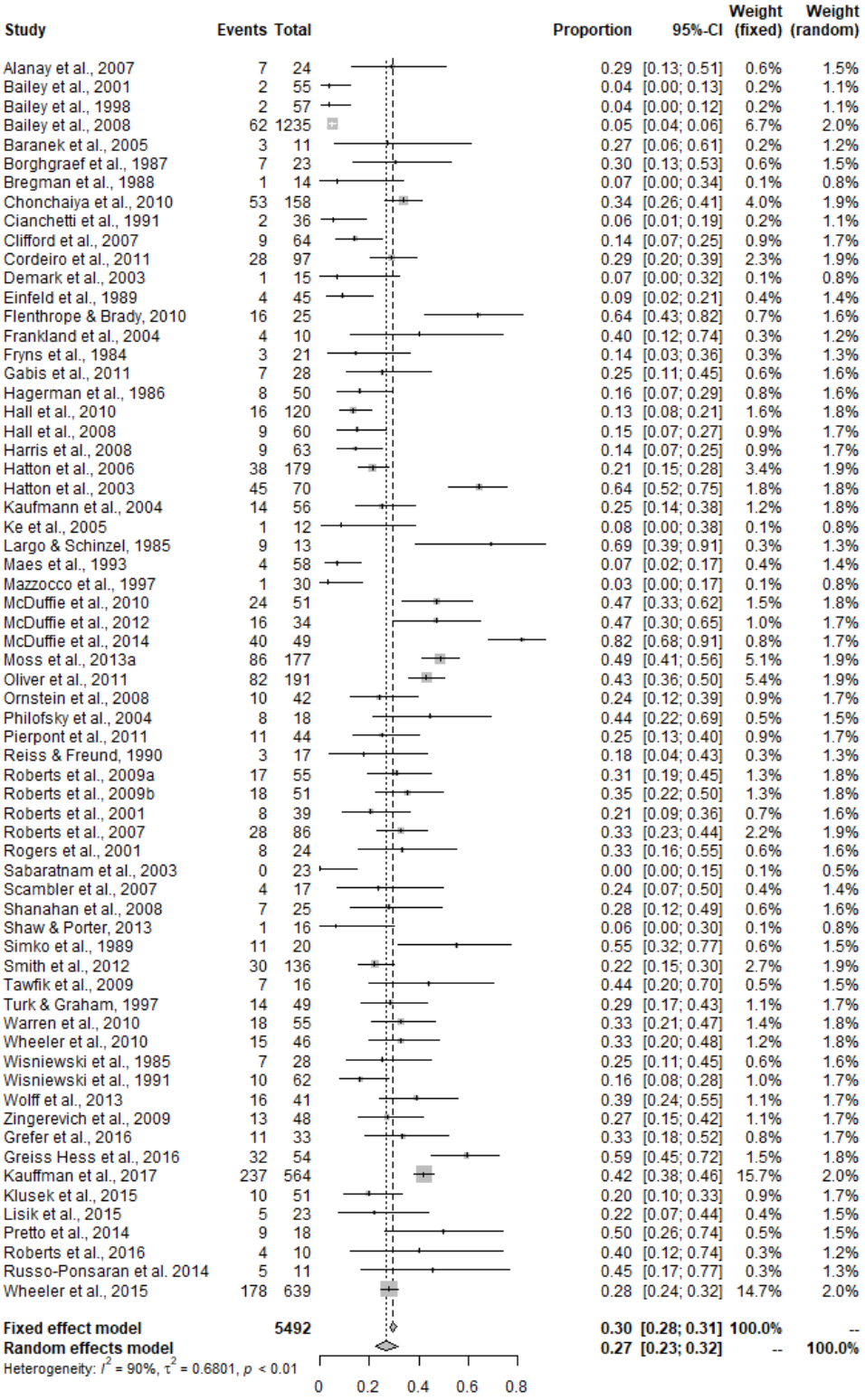

Figure 13. Pooled ASD prevalence estimates in Fragile X syndrome (random-effects model)

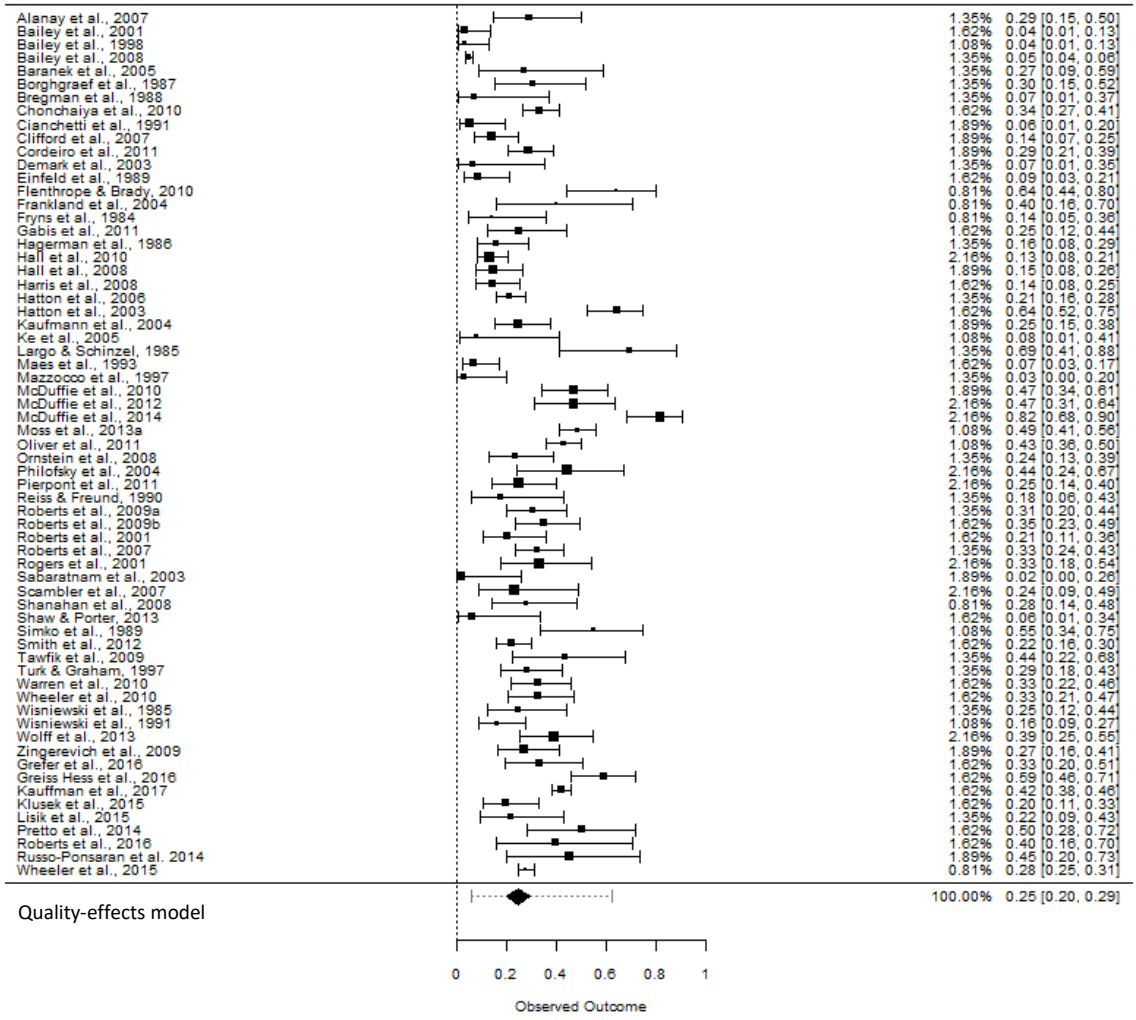

Figure 14. Pooled ASD prevalence estimates in Fragile X syndrome (quality-effects model)

Noonan syndrome

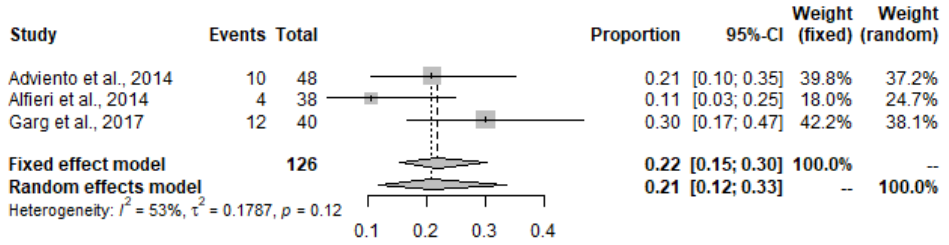

Figure 15. Pooled ASD prevalence estimates in Noonan syndrome (random-effects model)

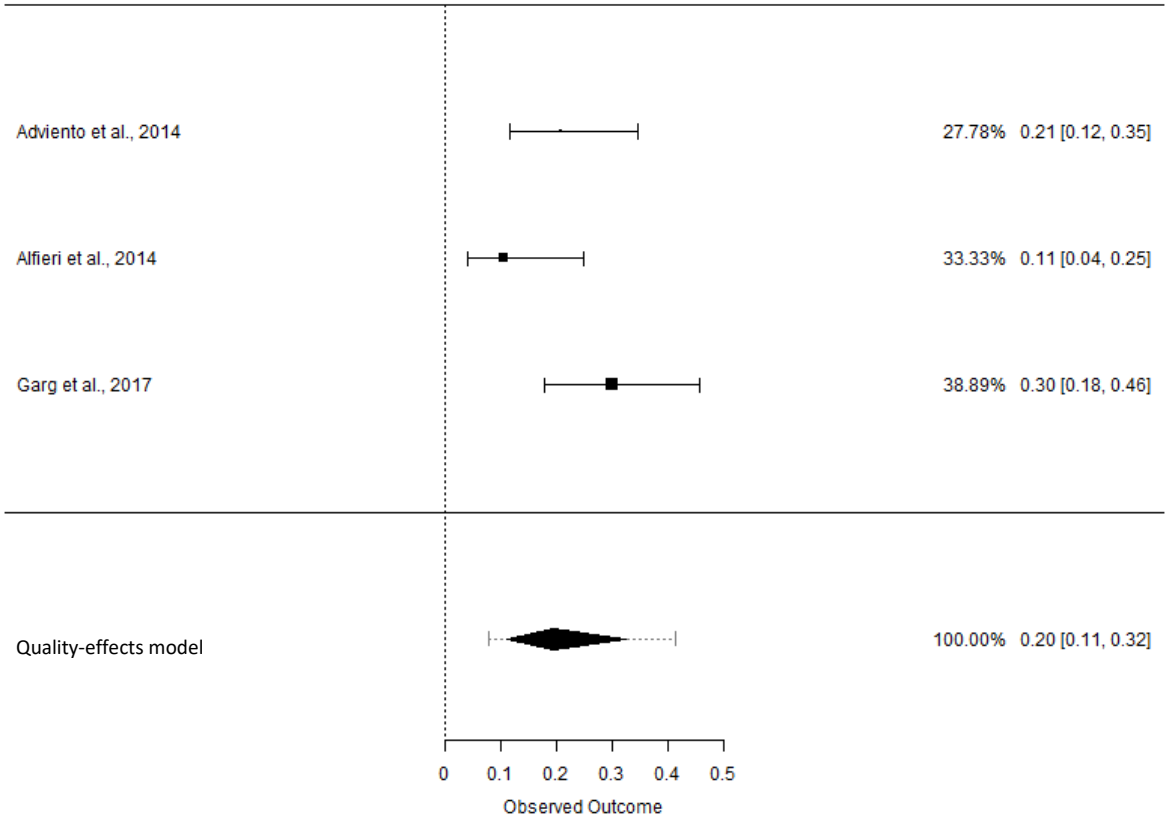

Figure 16. Pooled ASD prevalence estimates in Noonan syndrome (quality-effects model)

Neurofibromatosis

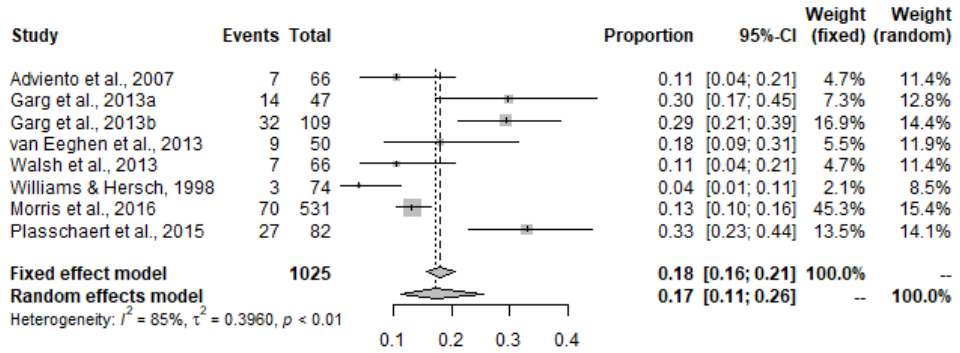

Figure 17. Pooled ASD prevalence estimates in Neurofibromatosis (random-effects model)

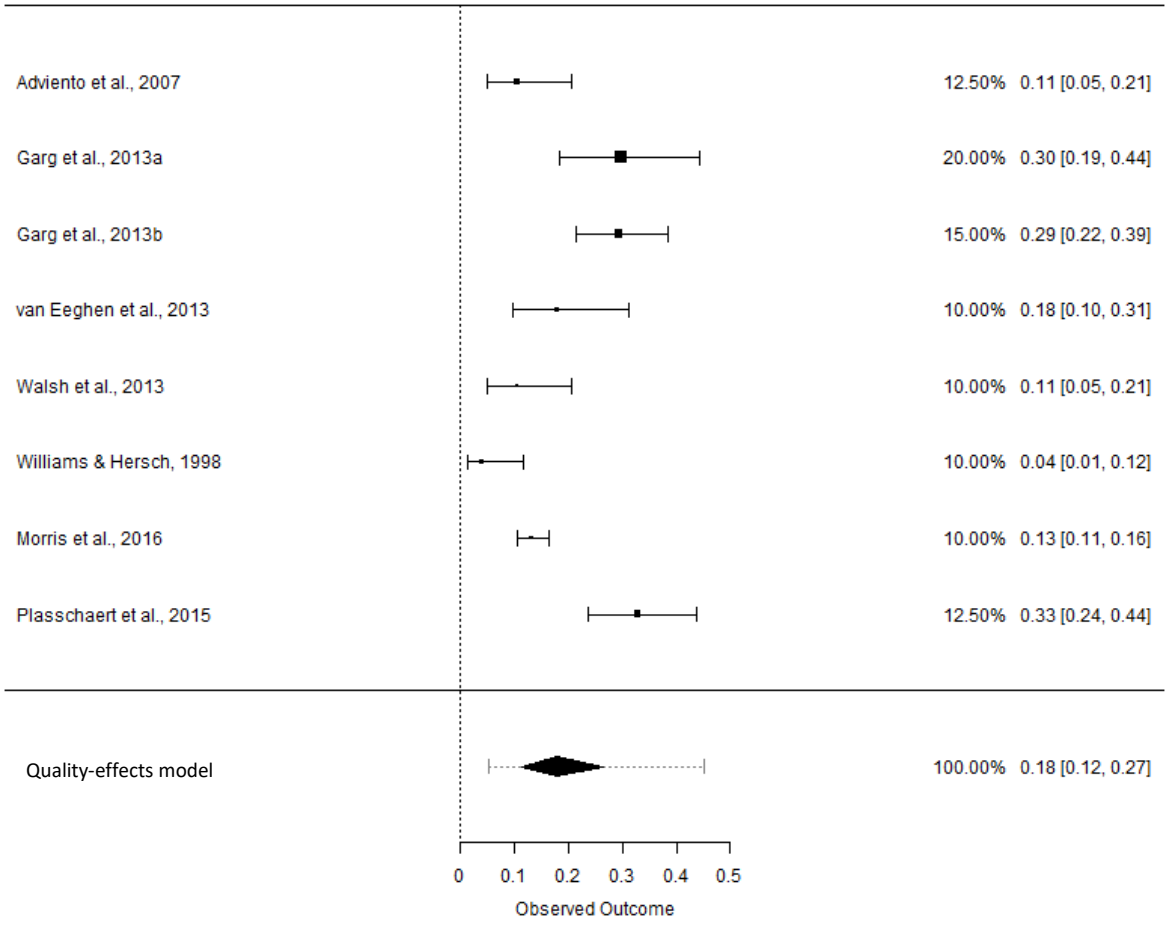

Figure 18. Pooled ASD prevalence estimates in Neurofibromatosis (quality-effects model)

Down syndrome

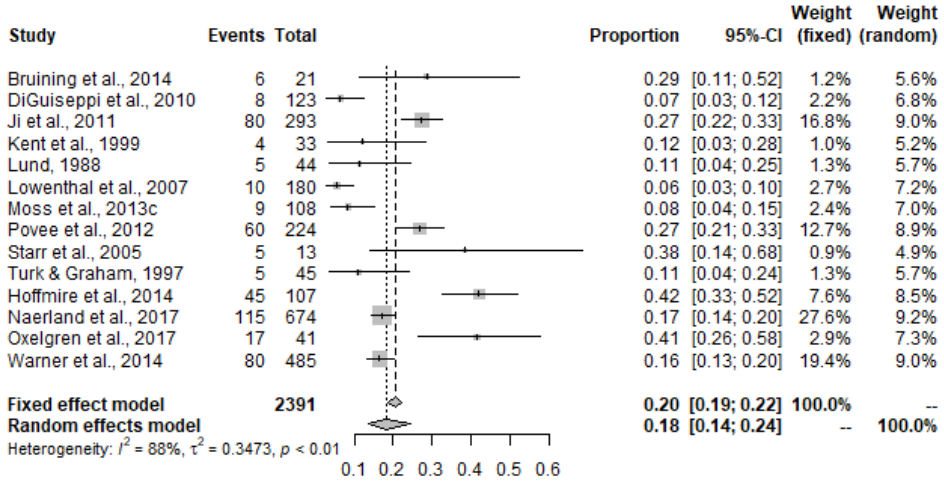

Figure 19. Pooled ASD prevalence estimates in Down syndrome (random-effects model)

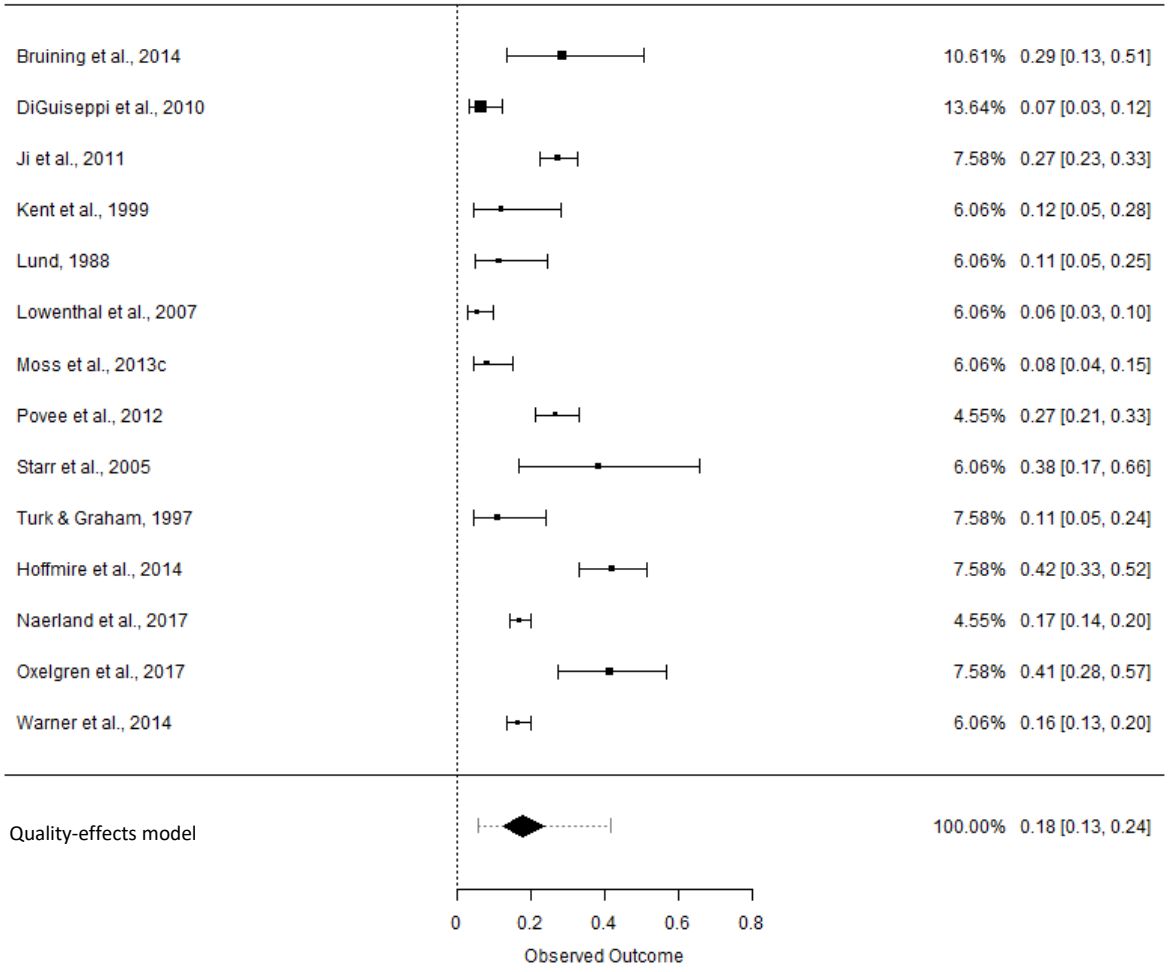

Figure 20. Pooled ASD prevalence estimates in Down syndrome (quality-effects model)

22q11.2 deletion syndrome (VCF)

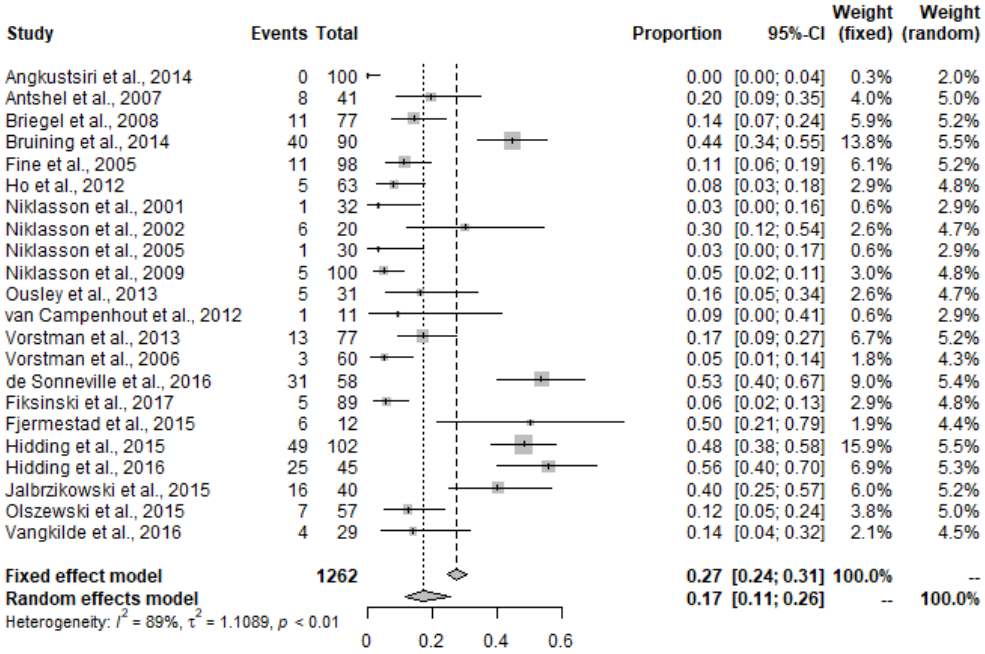

Figure 21. Pooled ASD prevalence estimates in 22q11.2 deletion syndrome (random-effects model)

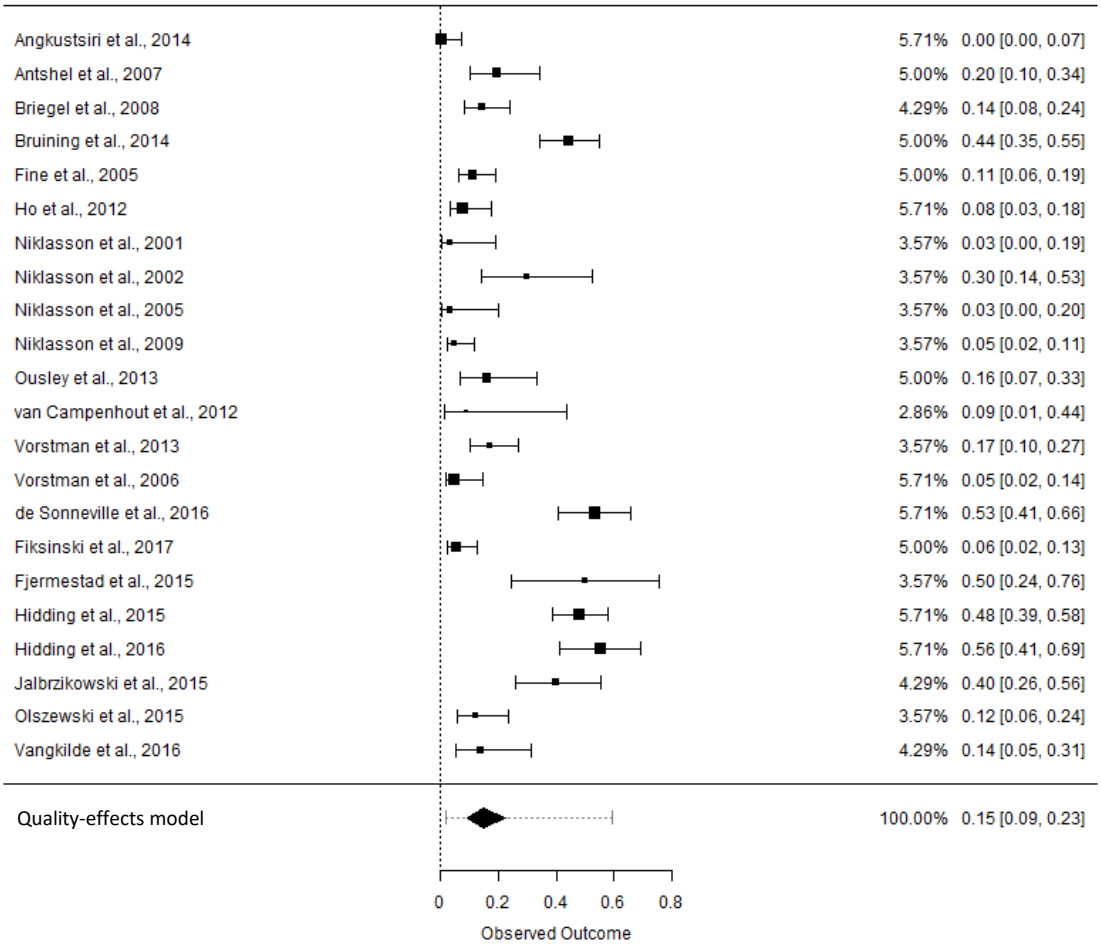

Figure 22. Pooled ASD prevalence estimates in 22q11.2 deletion syndrome (quality-effects model)

Phenylketonuria

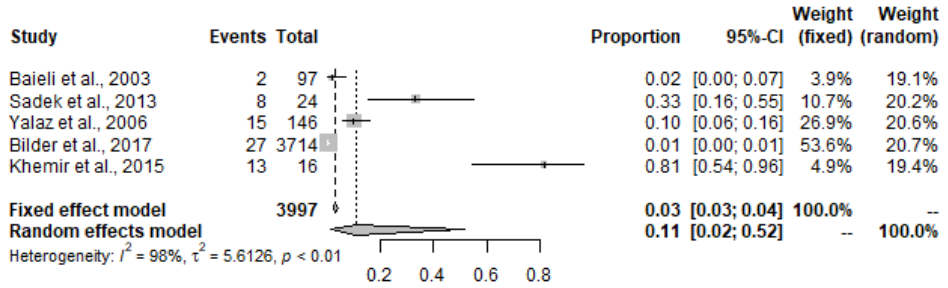

Figure 23. Pooled ASD prevalence estimates in Phenylketonuria (random-effects model)

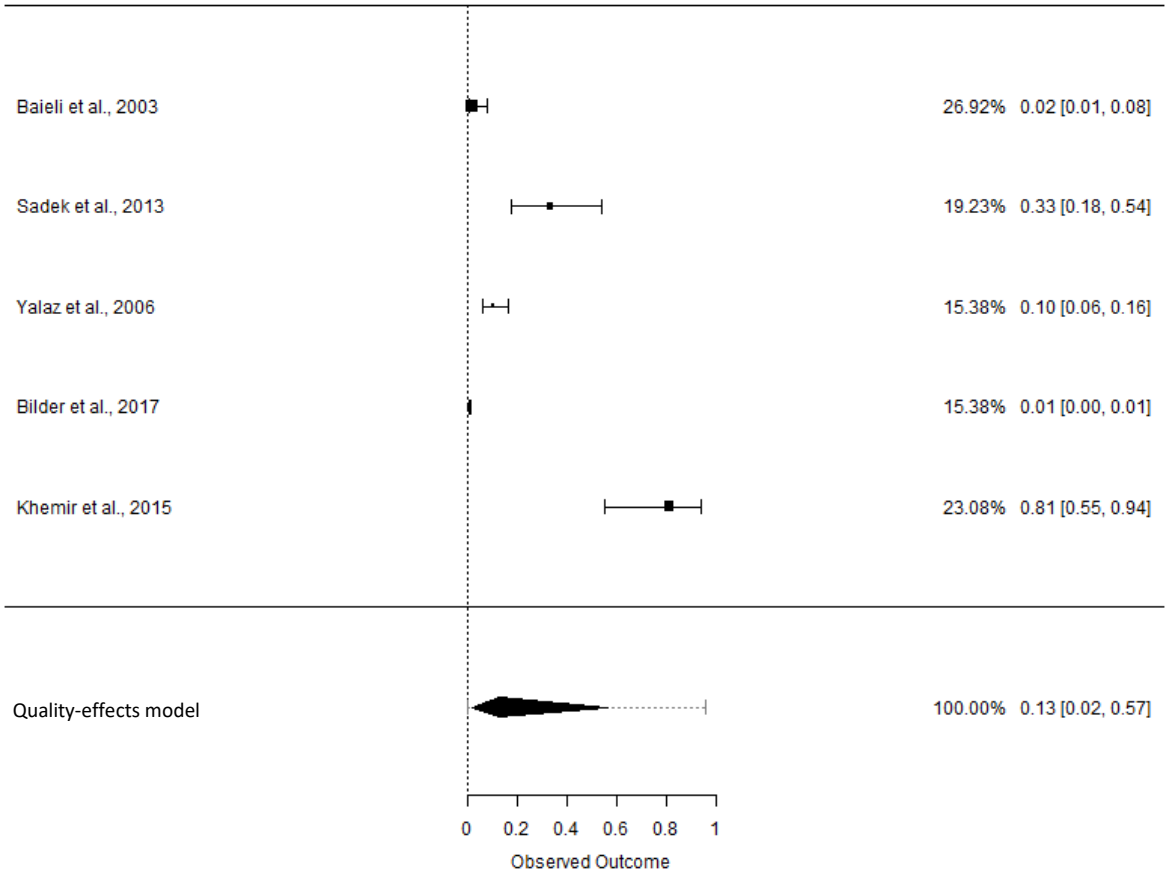

Figure 24. Pooled ASD prevalence estimates in Phenylketonuria (quality-effects model)

Williams syndrome

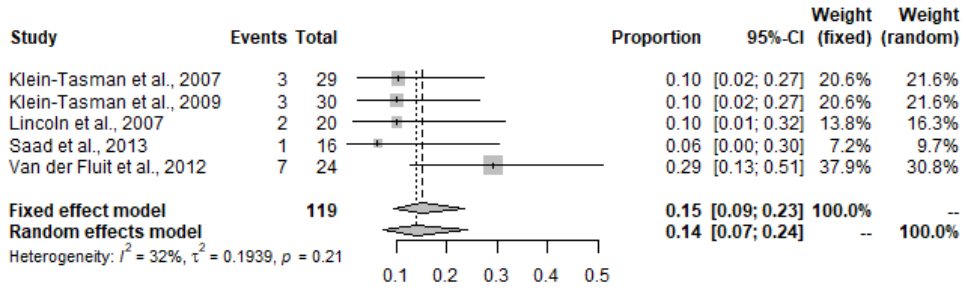

Figure 25. Pooled ASD prevalence estimates in Williams syndrome (random-effects model)

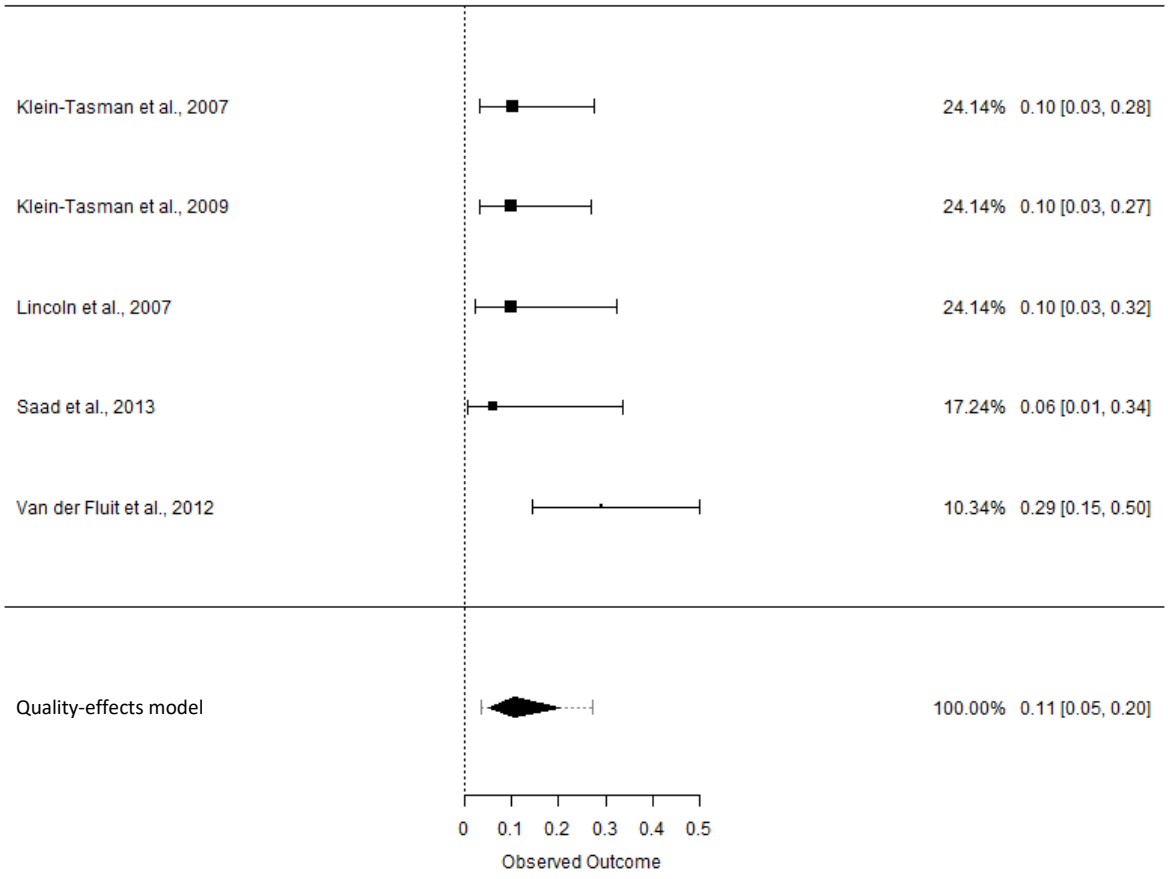

Figure 26. Pooled ASD prevalence estimates in Williams syndrome (quality-effects model)

Joubert syndrome

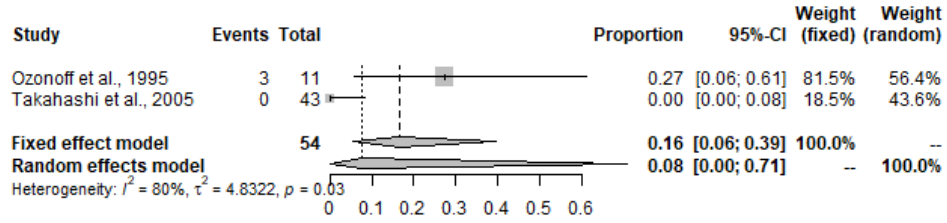

Figure 27. Pooled ASD prevalence estimates in Joubert syndrome (random-effects model)

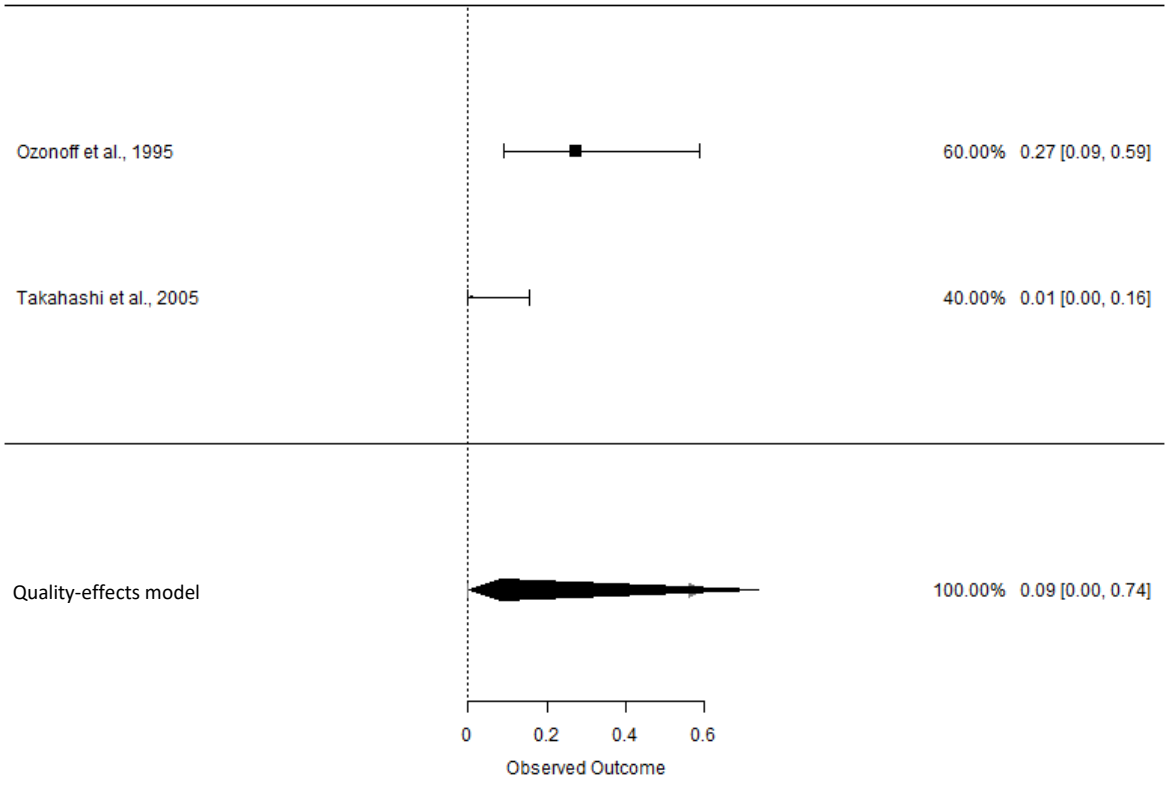

Figure 28. Pooled ASD prevalence estimates in Joubert syndrome (quality-effects model)

Moebius syndrome

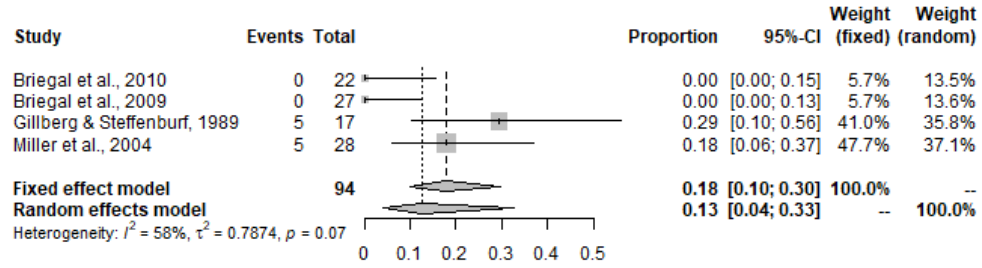

Figure 29. Pooled ASD prevalence estimates in Moebius syndrome (random-effects model)

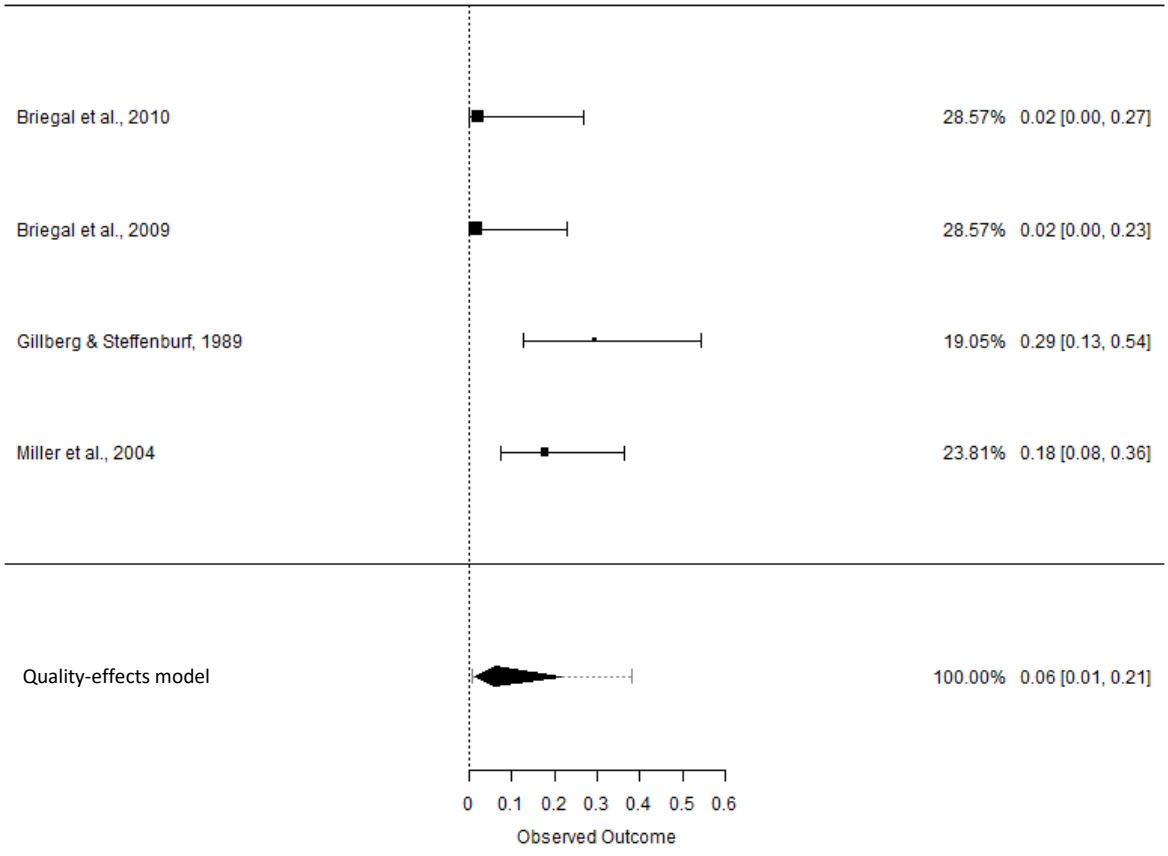

Figure 30. Pooled ASD prevalence estimates in Moebius syndrome (quality-effects model)

Appendix 4: Forest plots displaying pooled prevalence estimates grouped by quality rating for individual syndromes where a significant effect of quality criterion rating was observed

Angelman syndrome

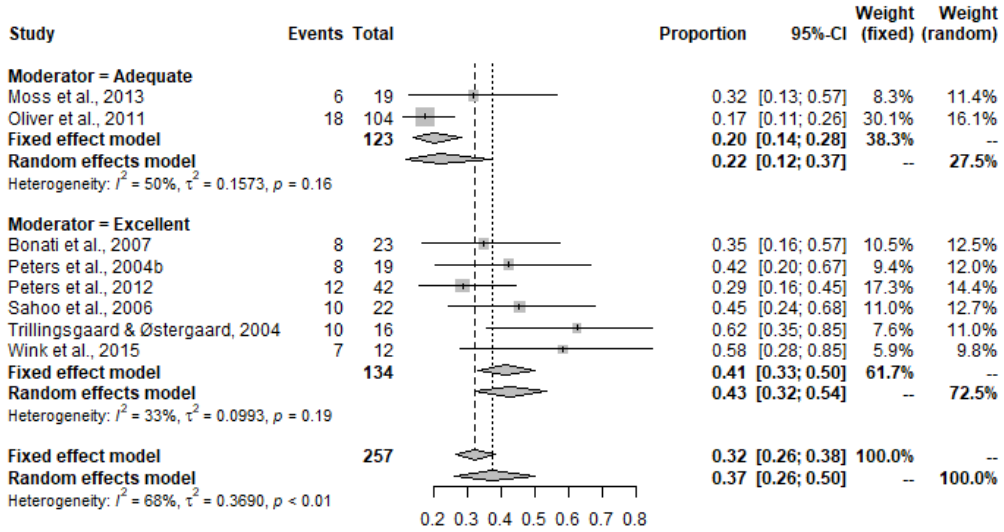

Figure 1. Pooled prevalence estimates for Angelman syndrome with studies grouped by Confirmation of Syndrome quality rating

Cornelia de Lange syndrome

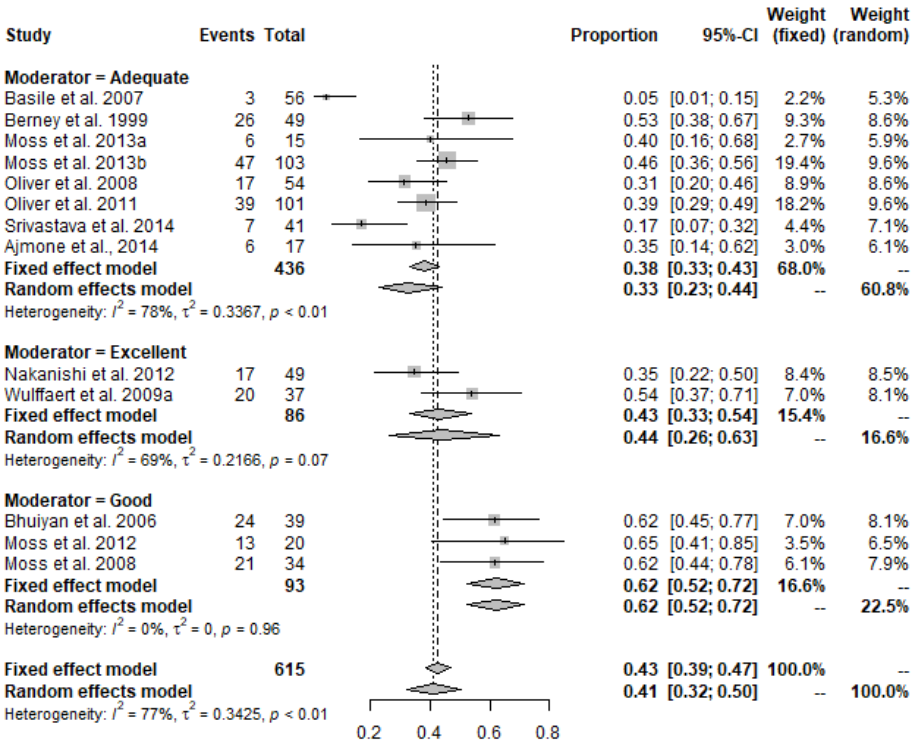

Figure 2. Pooled prevalence estimates for CdLS with studies grouped by ASD Assessment quality rating

Down syndrome

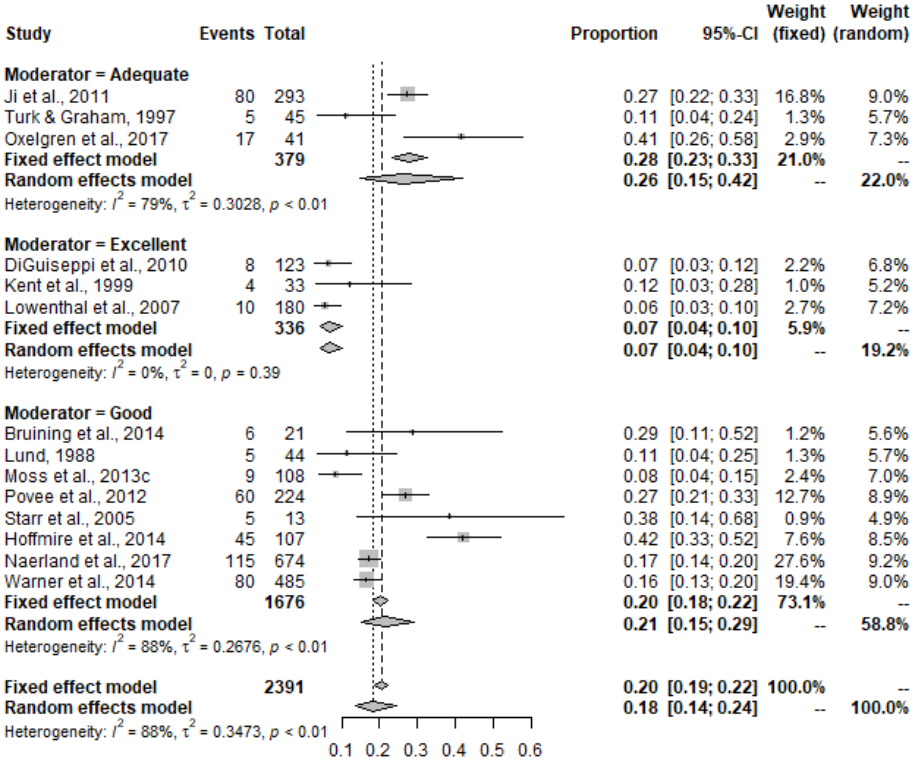

Figure 3. Pooled prevalence estimates for Down syndrome with studies grouped by Sample Identification quality rating

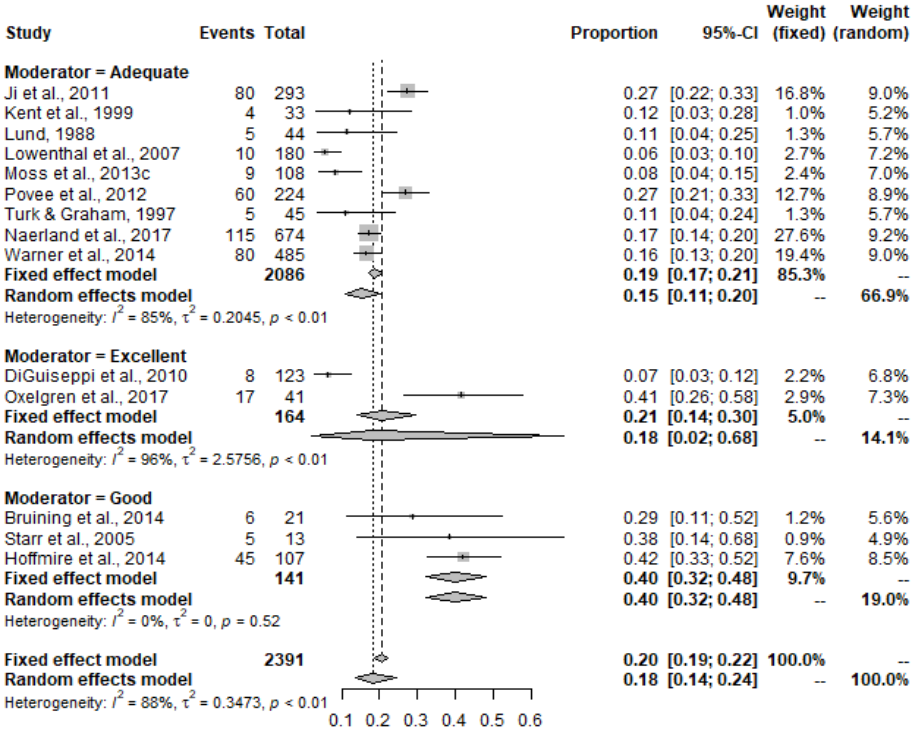

Figure 4. Pooled prevalence estimates for Down syndrome with studies grouped by ASD Assessment quality rating

Fragile X syndrome

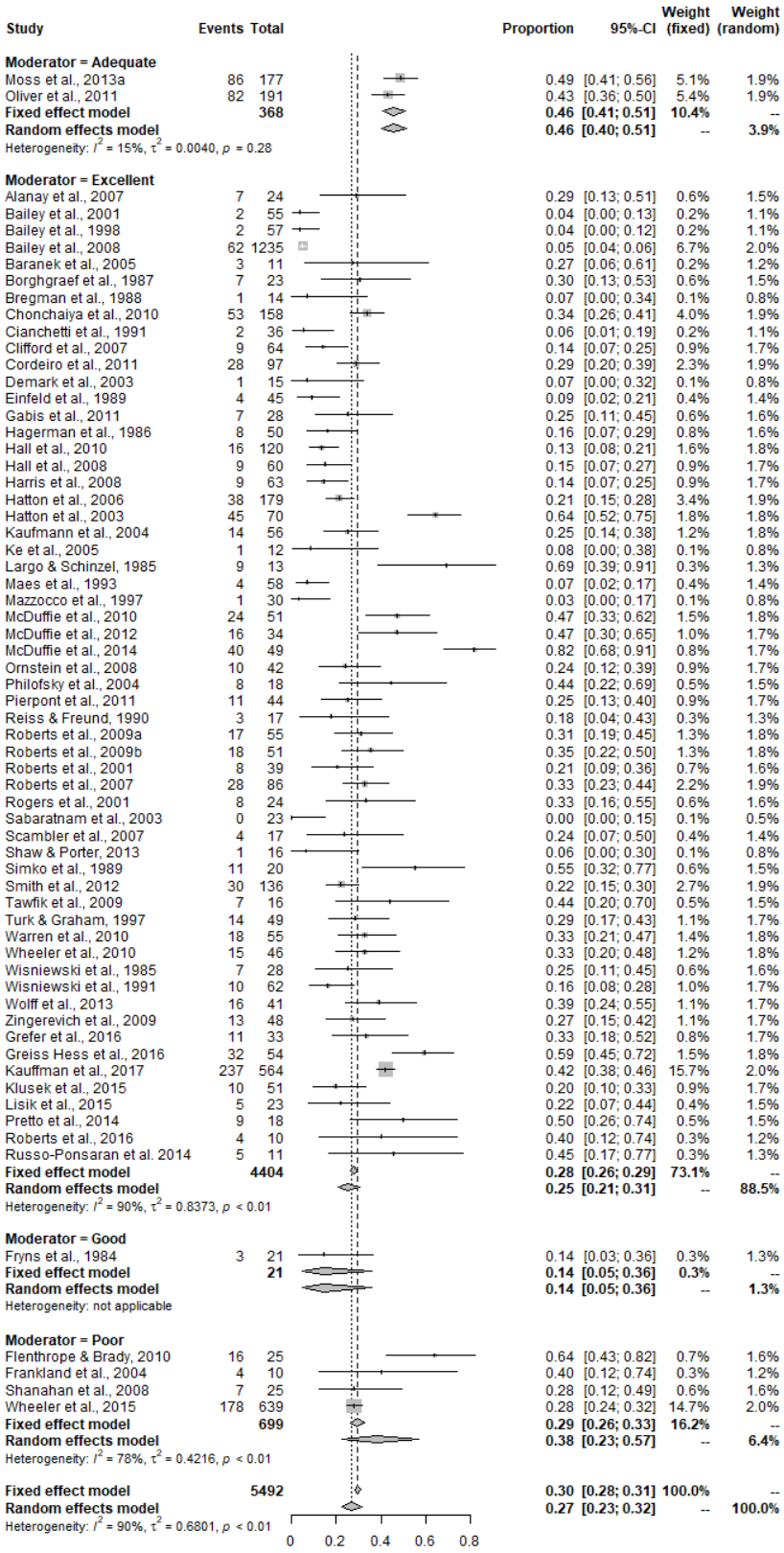

Figure 5. Pooled prevalence estimates for Fragile X syndrome with studies grouped by Confirmation of Syndrome quality rating

Joubert syndrome

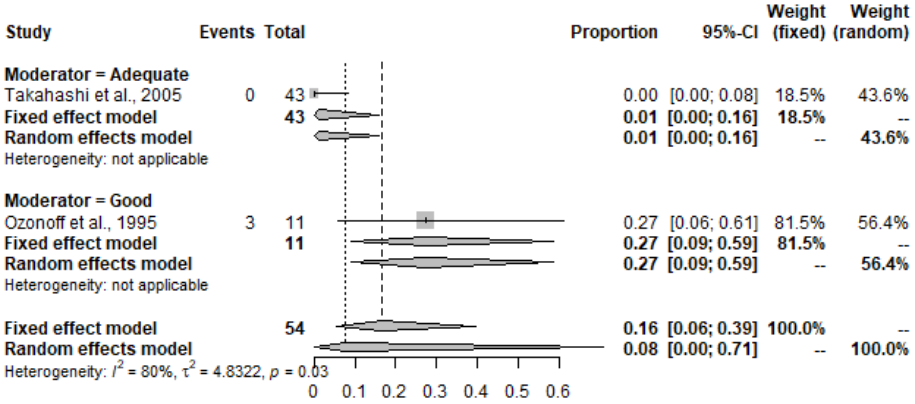

Figure 6. Pooled prevalence estimates for Joubert syndrome with studies grouped by Confirmation of Syndrome quality rating

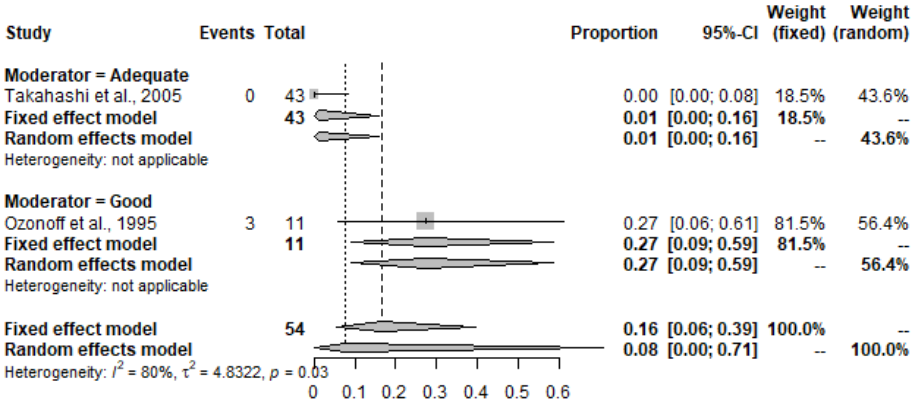

Figure 7. Pooled prevalence estimates for Joubert syndrome with studies grouped by ASD Assessment quality rating

Moebius syndrome

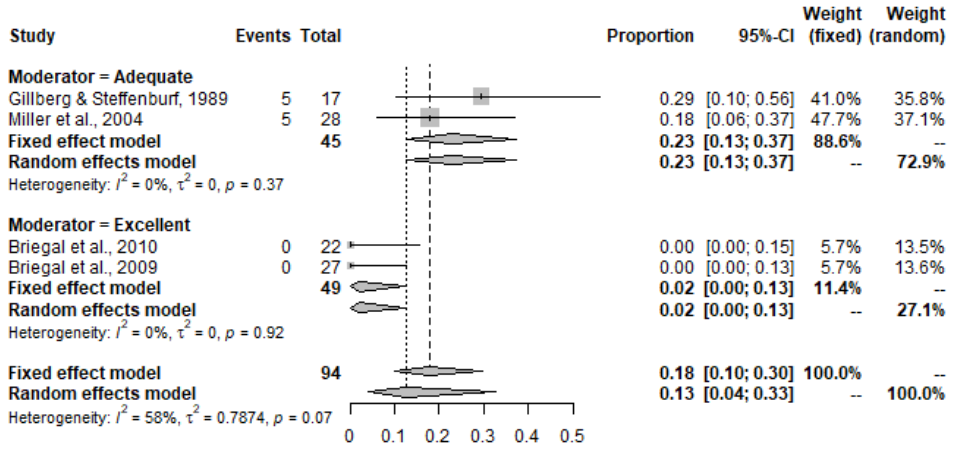

Figure 8. Pooled prevalence estimates for Moebius syndrome with studies grouped by ASD Assessment quality rating

Neurofibromatosis

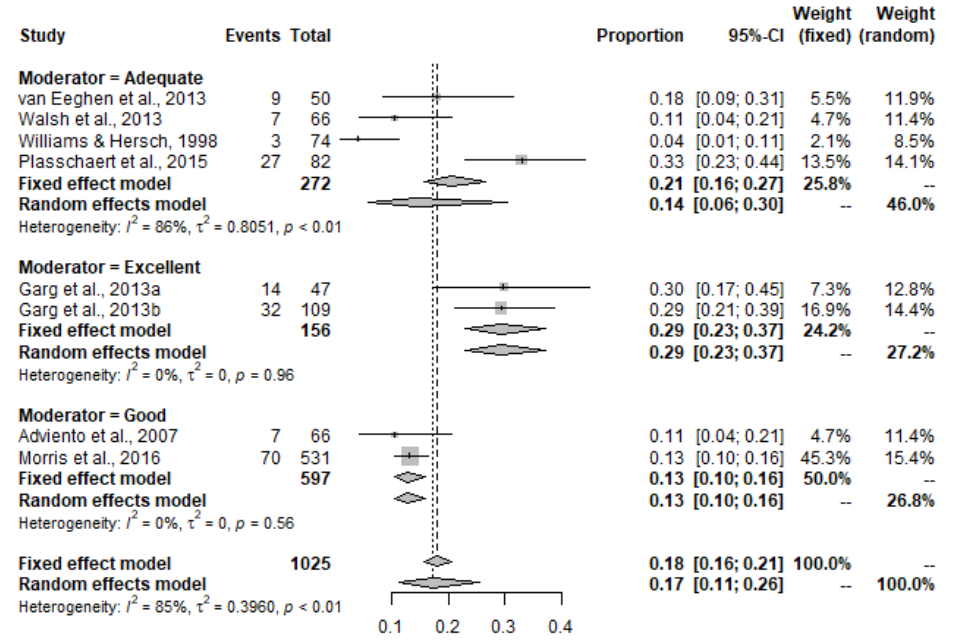

Figure 9. Pooled prevalence estimates for Neurofibromatosis with studies grouped by Sample Identification quality rating

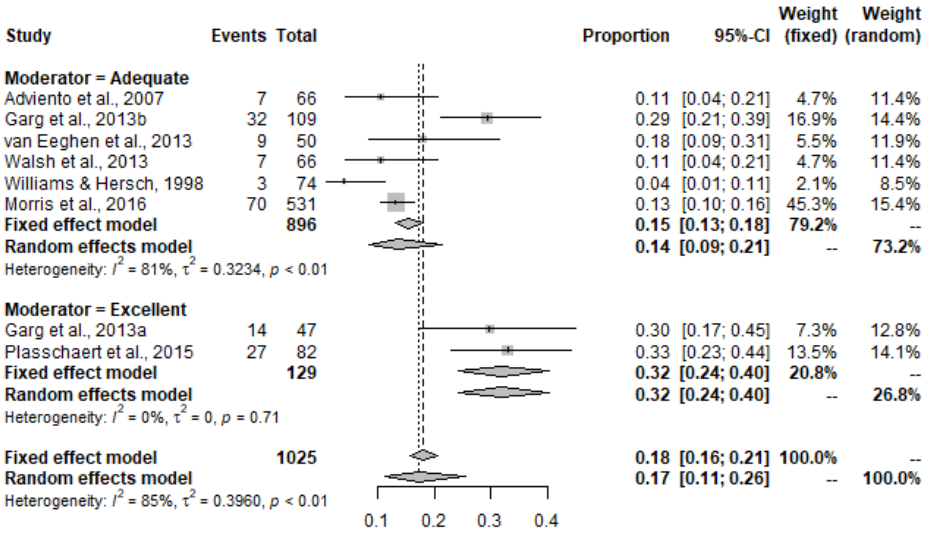

Figure 10. Pooled prevalence estimates for Neurofibromatosis with studies grouped by ASD Assessment quality rating

Phenylketonuria

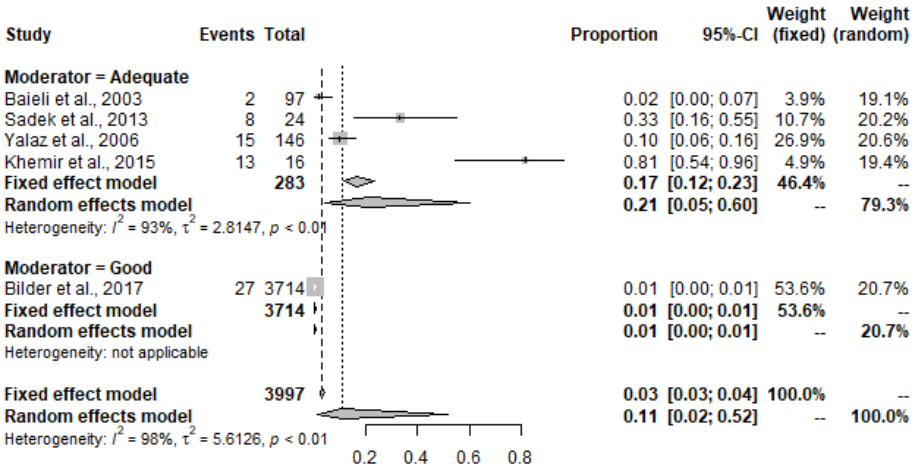

Figure 11. Pooled prevalence estimates for Phenylketonuria with studies grouped by Sample Identification quality rating

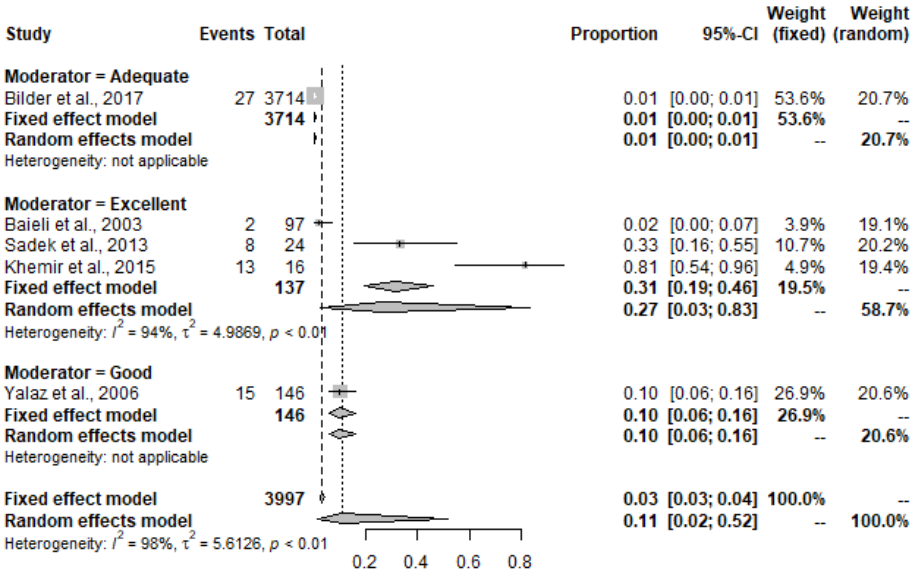

Figure 12. Pooled prevalence estimates for Phenylketonuria with studies grouped by Confirmation of Syndrome quality rating

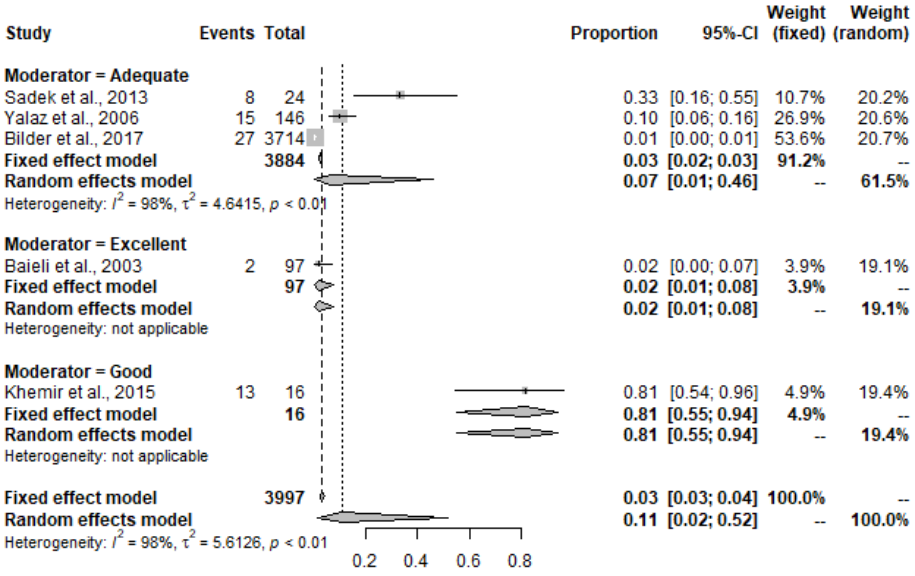

Figure 13. Pooled prevalence estimates for Phenylketonuria with studies grouped by ASD Assessment quality rating

Rett syndrome

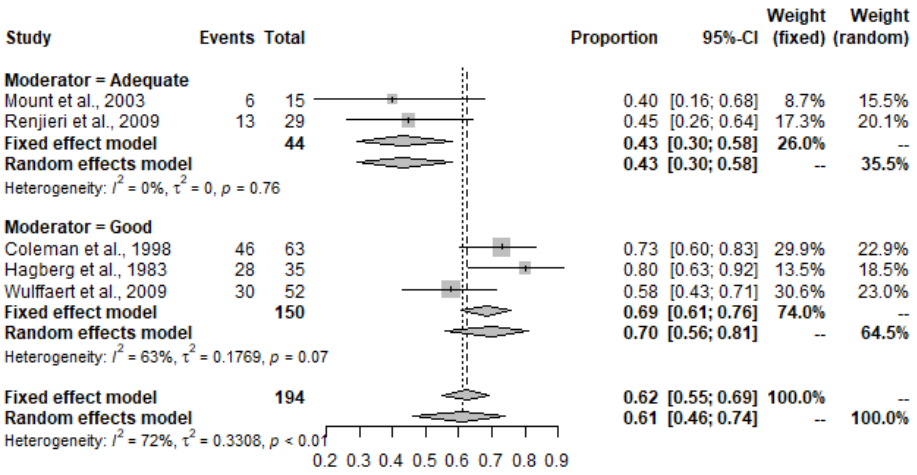

Figure 14. Pooled prevalence estimates for Rett syndrome with studies grouped by Sample Identification quality rating

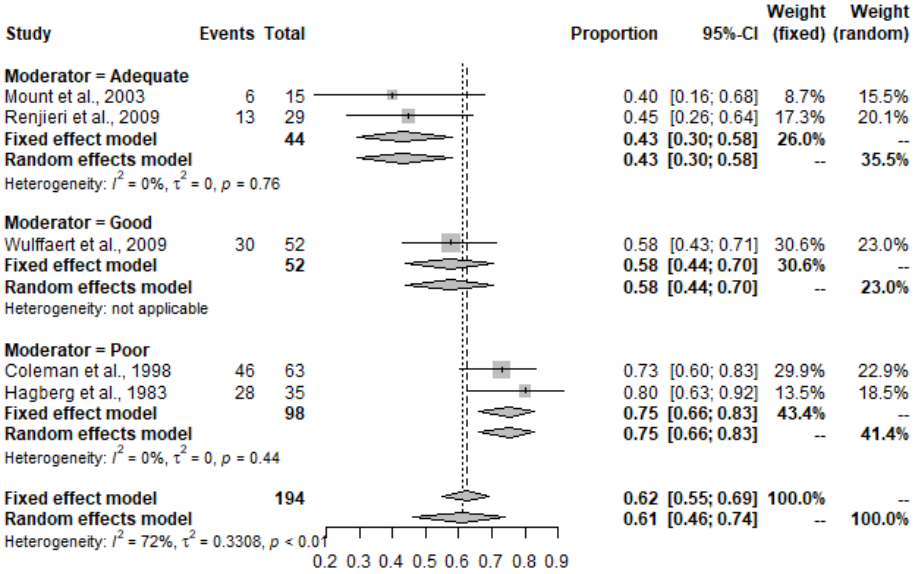

Figure 15. Pooled prevalence estimates for Rett syndrome with studies grouped by ASD Assessment quality rating

Sotos syndrome

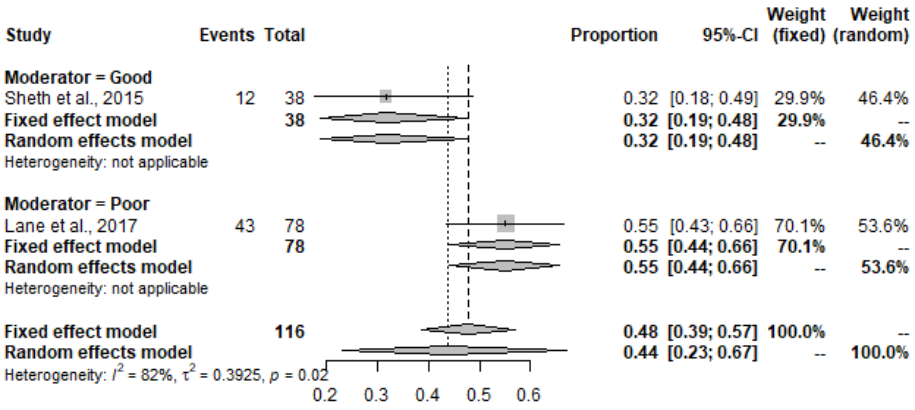

Figure 16. Pooled prevalence estimates for Sotos syndrome with studies grouped by Confirmation of Syndrome quality rating

Tuberous sclerosis complex

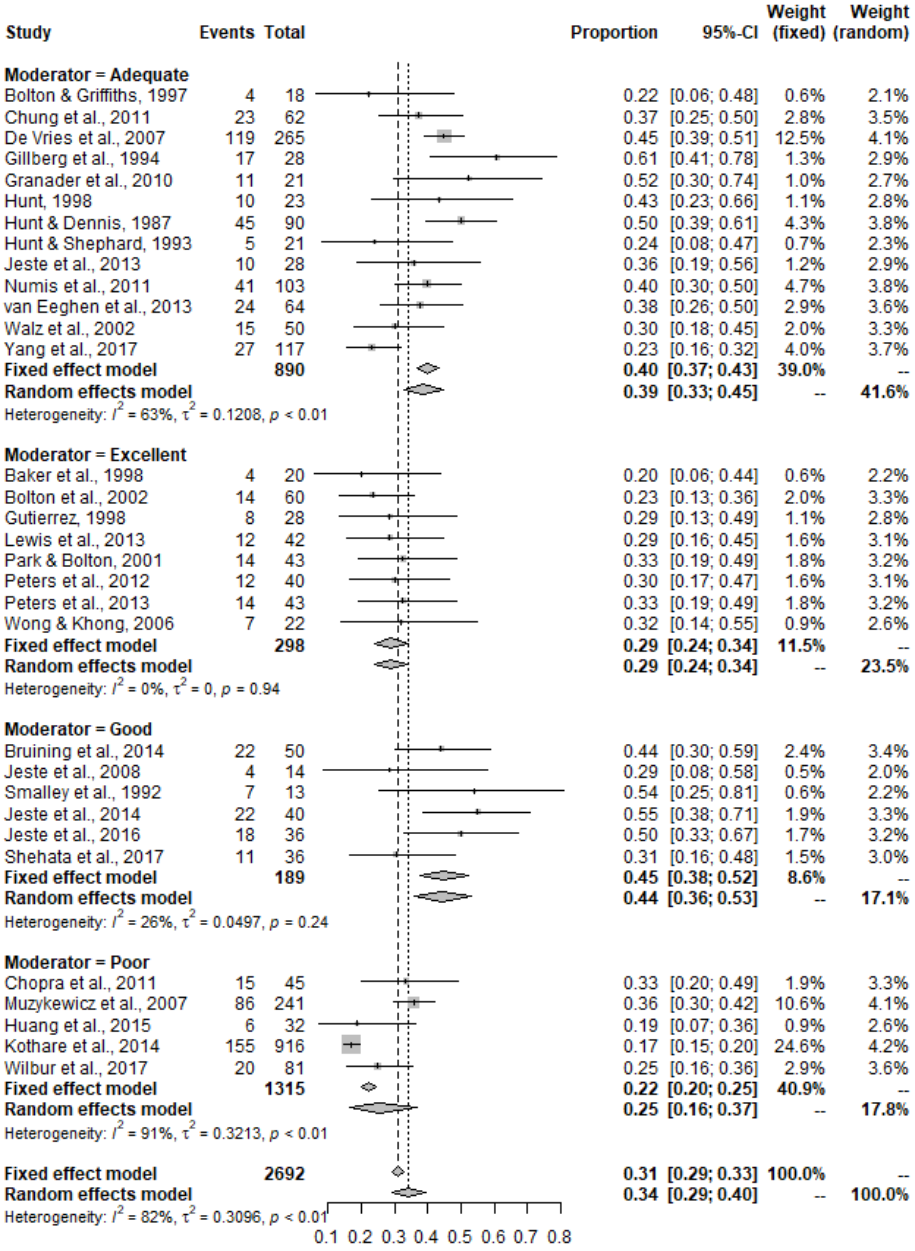

Figure 17. Pooled prevalence estimates for tuberous sclerosis complex with studies grouped by ASD Assessment quality rating

22q11.2 deletion syndrome

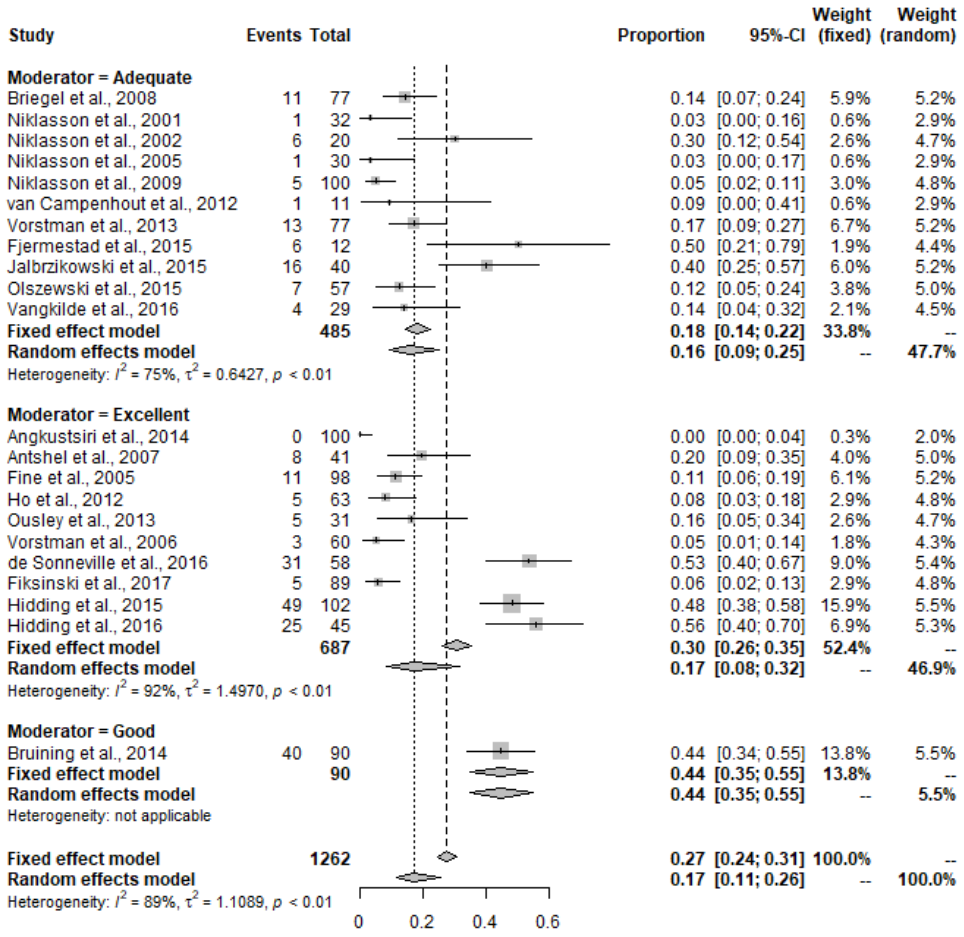

Figure 18. Pooled prevalence estimates for 22q11.2 deletion syndrome with studies grouped by ASD Assessment quality rating

Williams syndrome

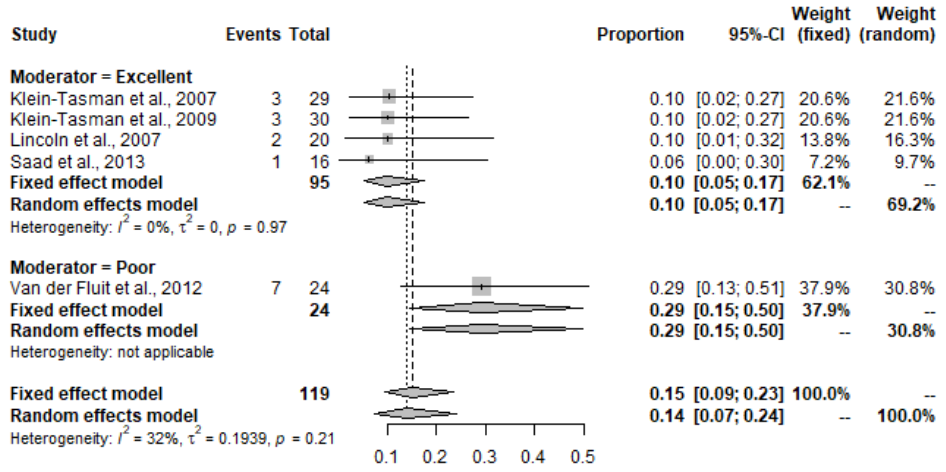

Figure 19. Pooled prevalence estimates for Williams syndrome with studies grouped by Confirmation of Syndrome quality rating

Appendix 5: Results of a multiple predictor meta-regression

The proportion of male participants, mean age and mean IQ were entered into a multiple predictor meta-regression. Mean IQ, rather than the proportion of the sample with an intellectual disability, was selected as a measure of intellectual ability as this resulted in a larger number of studies reporting all participant characteristics (*n*=45). In this multiple meta-regression model, only mean IQ significantly predicted ASD prevalence estimates (*p*=.017), evidencing a negative association between IQ and prevalence estimates. The influence of gender and mean age were both non-significant (p>.05). From the results of this regression model, prevalence estimates for each syndrome were calculated controlling for the predictive effect of gender, age and IQ. These are presented in Table 1. For Angelman syndrome, Cornelia de Lange syndrome, CHARGE syndrome, Joubert syndrome, Rett syndrome and Sotos syndrome no studies provided all participant characteristics, therefore adjusted prevalence estimates were not available.

Table 1. Pooled prevalence estimates controlling for the predictive effect of gender, age and IQ

	Pooled prevalence estimate, controlling for the influence of gender, age and IQ		
	Prevalence estimate	*Lower confidence interval*	*Upper confidence interval*
22q11.2 deletion syndrome	0.76	0.60	0.87
Cohen syndrome	0.76	0.64	0.85
Down syndrome	0.70	0.56	0.82
Fragile X syndrome	0.74	0.58	0.85
Moebius syndrome	0.75	0.53	0.89
Neurofibromatosis	0.80	0.59	0.92
Noonan syndrome	0.79	0.38	0.91
Phenylketonuria	0.71	0.51	0.84
Tuberous sclerosis complex	0.78	0.51	0.88
Williams syndrome	0.70	0.53	0.82

Appendix 6: Letter to schools

Document redacted

Appendix 7: Study information sheet

Document redacted

Document redacted

Document redacted

Document redacted

Document redacted

Appendix 8: Consent forms

Document redacted

Document redacted

Document redacted

Document redacted

Document redacted

Document redacted

Document redacted

Document redacted

Appendix 9: Early Social Cognition Scale administration and scoring (from Powis, 2014)

1.1 Administration instructions

Helping

Control trials

Following a short warm up period infants were sat at a table opposite Experimenter one (E1). For the 'pen' condition, E1 was seen to use a pen for drawing. The experimenter then stopped drawing, put the lid on the pen, intentionally threw it on the floor and did not reach for it. For the 'paper balls' condition, E1 set up three paper balls on her side of the table and three paper balls on the infant's side of the table. The experimenter then sat back down and picked up each of the paper balls 'one-by-one' using a pair of tongs and then placed them back on the table.

Experimental trials

For the 'pen' condition, Experimenter two (E2) was seen to use a pen for drawing, she then 'accidentally' dropped the pen on the floor and unsuccessfully reached for it. For the 'paper balls' condition, E2 picked up each paper ball with a pair of tongs and placed them in a cardboard container. She then attempted to reach for the three paper balls on the infant's side of the table but failed because they were too far away.

Coding

In each trial, infants' behaviour was coded for whether or not they performed the required *target* behaviour. For the pen trial this involved the infant passing the pen back to the experimenter. For the paper balls trial, it involved the infant passing or pushing the paper balls towards the experimenter.

Pass/fail criterion and rationale

Infants were coded as having 'passed' the helping task if they successfully demonstrated one of the target behaviours. This was considered appropriate as unlike some of the other

tasks included in the scale, 'helping' behaviour was considered unlikely to occur by chance. Therefore one demonstration was deemed sufficient to indicate that the infant had acquired the skill and passed the task.

Re-enactment of intended acts

For this task infants were sat at a table opposite the experimenter. Three experimental trials were administered to each infant and all followed the same procedure. For each trial the experimenter presented the infant with an object pair that could be used to perform a target act – a loop that could be hung over a protruding peg, some beads that could be dropped into a cup, or a square with a hole in it that could be stacked upon a protruding peg. For each trial the experimenter modelled the intention to perform the target act but ultimately failed to perform the act.

For the loop and peg, the experimenter picked up the loop and moved her hand towards the peg but released it inappropriately each time so that instead of hanging over the peg, the loop 'accidentally' fell to the table. Initially the loop was released slightly too far to the left, then too far to the right and then too low.

For the beads and cup, the experimenter picked up the beads and attempted to drop them into the cup but released them inappropriately each time so that they 'accidentally' fell to the table instead. Initially the beads were lowered just so that they touched the lip of the cup but then released so that they fell to the side. On the next attempt the beads were suspended too far in front of the cup and so fell to the table when released. On the final attempt the experimenter gathered the beads loosely in her hand but then scraped her hand over the top of the cup so that the beads fell to the side rather than inside the cup.

For the square and protruding peg the experimenter picked up the square and attempted to place it upon the peg, however each time the experimenter failed to align it correctly so that it 'accidentally' overshot the peg. Initially the square overshot to the right, then to the left and finally to the front.

After the experimenter had demonstrated the three failed attempts she offered the object pair to the infant. During the experimental procedure the experimenter did not provide the infant with any prompts or cues, however, the experimenter gained the infant's attention by saying "Oh, look what I have here", "What's this?", and "Now it's your turn".

Coding

For each trial infants' behaviour was coded for whether or not they went on to perform the target act themselves: for the loop and peg this involved them hanging the loop over the protruding peg; for the beads and cup this involved them dropping the beads inside the cup; and for the square and peg this involved them stacking the square over the protruding peg.

Pass/fail criterion

Infants were coded as having passed the task if they successfully performed the target act in *two or more of the three trials.* This pass/fail criterion was deemed necessary to reduce the possibility that infants might 'pass' one trial simply by chance or because the apparatus 'afforded' a particular response from that infant. It was decided that two or more target acts were less likely to occur 'just by chance' and therefore it was deemed that this provided sufficient evidence that the infant possessed the social cognitive skill.

Gestures (Point and Gaze)

Due to the identical experimental procedures used to assess these communicative cues, trials for both cues were administered together in one procedure.

Warm up phase

Before the task began each infant took part in a warm up phase. This was to familiarise them with the hiding procedure and the containers used. Infants were sat at a table next to E2 and across a table from E1. E1 placed a pair of open containers in front of the infant and then brought out a small toy. E1 then announced "Look, I'll hide it". As the infant watched, E1 placed the toy in one of the containers and then placed the lids on both. E2

then encouraged the infant to retrieve the toy by saying “Where’s the toy” and “Can you get the toy?” This warm up hiding procedure was repeated three times with three different sets of containers.

Control trials

Each infant participated in four control trials. These control trials were administered to check that search performance was indicative of understanding the experimenter’s intentions and not simply due to low level attentional cueing. For each control trial E1 placed a pair of open containers on the table then produced a small toy. If the infant showed interest in the toy E1 then placed a movable screen in front of the two containers, lowered the toy behind the screen and said “Now I’ll hide it”. At this point E1 then quickly pushed the containers together, hid the toy in one, and then moved them apart again. The distance between each container ensured that the infant could not grab both containers at the same time. Following the hiding procedure E1 removed the screen and gave one of two non-communicative control cues:

- *Control Point* – E1 performed a ‘distracted point’ by holding out her hand and slightly extending her index finger. E1 simply looked down at her hand with an expression that indicated she was preoccupied by something on her hand.
- *Control gaze* – E1 gazed at the container with an absent minded facial expression, eyes unfocused with a neutral facial expression.

Following each non communicative cue E2 then encouraged the infant to retrieve the toy by saying “Where’s the toy?” and “Can you get the toy?” Each infant received two control gaze trails and two control point trials, which were represented in one of four different counterbalanced orders.

Experimental trials

As before, for each trial E1 placed a pair of containers in front of the infant and produced a toy. If the infant showed interest in the toy E1 then placed a moveable screen in front of the two containers, lowered the toy behind the screen and said “Now I’ll hide it”. During this hiding procedure E2 showed the infant that she was watching by alternating her gaze between the containers and the infant then announcing “I can see”. After hiding had been

completed E1 pushed the containers apart and removed the screen. E1 then turned away from the table in order to place the screen behind her. At this point, while E1 was not looking, E2 established eye contact with the infant and gave one of two communicative gestures:

- *Point* – E2 extended her index finger and pointed at the container expressing intent through raised eyebrows.
- *Ostensive gaze* – E2 gazed at the target container and then back to the infant expressing intent through raised eyebrows.

Following each communicative cue E1 then turned back to the table and encouraged the infant to retrieve the toy by saying "Where's the toy" and "Can you get the toy?". Each infant received two gaze trials and two point trials which were presented in one of four different counterbalanced orders.

To minimise the possibility of perseveration errors being made, each pair of containers were different in colour and shape and the same pair were never used on successive trials. For each trial, it the infant attempted to open a container but could not quite manage to, then one of the experimenters assisted the infant. Furthermore, if the infant chose the incorrect box and did not find the toy, then the experimenters opened the correct box and gave the toy to the infant. This was done to ensure that the infant did not become frustrated and disengage from the task.

Coding

For each trial the box that the infant first selected and attempted to open was recorded. If the infant selected the container that the toy was hidden in this was coded as correct. If the infant selected the container without the toy this was coded as incorrect.

Pass/fail criterion

Infants were coded as having passed the point trials if they successfully chose both of the correct containers following each point gesture. Similarly, infants were coded as having passed the gaze trials if they successfully chose both the correct containers following each gaze gesture. This pass/fail criterion was deemed necessary to reduce the possibility that

infants might 'pass' the task simply be selecting the correct location by chance. It was considered much less likely that infants would select the correct location by chance on *two* consecutive experimental trials.

Cooperation (Tubes-with-handles and Trampoline)

Cooperation: Tubes-with handles

In the *Tubes-with-handles* task the infant's goal was to retrieve a toy that had been hidden inside a tube. This tube could be pulled apart by pulling the handles at each end of the tube. However, the length of the tube made it impossible for the infant to perform this goal alone and therefore to be successful the infant was required to 'work together' and cooperate with the experimenter. For each infant, the experimental procedure included a number of steps.

- *Familiarisation and demonstration*: The task began with a brief familiarisation period in which the infant was shown the tube and encouraged to hold each of the handles. After the infant was familiar with the apparatus, E1 and E2 pulled the tube apart and E1 placed an attractive object inside. The two experimenters pushed the tube back together and placed it on the floor. E1 and E2 then proceeded to demonstrate how the toy could be retrieved by each of them pulling the handles at each end. Following this demonstration E1 produced another attractive object and placed it inside the tube as before, then pushed it back together with E2.
- *Experimental trial one*: Following the familiarisation and demonstration period E1 invited the infant's participation by alternating gaze between the infant and the tube. If the infant was immediately successful and cooperated with the experimenter to open the tube then trial 2 was administered. However, if the infant was not successful within 30 seconds E1 and E2 carried out the demonstration phase again. Following the second demonstration the infant was encouraged to participate again, this time using verbal cues such as "Come and try" and "Look, tube!". If the infant was still unsuccessful the demonstration was repeated for a third time but this time E2 also encouraged the infant to stand by her and hold the handle together with her. If the infant was unsuccessful after three demonstrations the task was discontinued.

- *Experimental trial two*: Following trial one, E1 produced another toy and placed it inside the tubes as before. The infant's participation was then prompted by E1. Once object was retrieved trial three was administered. If the infant was not successful after 60 seconds then the task was discontinued.
- *Experimental trial three*: Following trial two, E1 produced another toy and placed it inside the tube, as before. The infant's participation was once again prompted by E1. However, in this trial, unlike trials one and two, when the infant picked up their side of the tube E1 dropped her side of the tube and placed her hands and face downwards for an interruption period of 15 seconds. Following the interruption period E1 looked back up, picked up the tube and continued as before. If the infant had disengaged E1 prompted the infant's participation to continue. After the infant had retrieved the object trial four was administered.
- *Experimental trial four*: The procedure for trial three was repeated.

Cooperation: Trampoline

The procedure for the *Trampoline* task was very similar to the *Tubes-with-handles* task. However, in the task the infant's goal was to bounce a toy up and down on a handheld trampoline. Importantly, due to joints on the side of the trampoline if two people did not hold it at the same time it would collapse. Therefore to be successful on this task the infant was required to 'work together' and cooperate with the experimenter. For each infant the experimental procedure included a number of steps.

- *Familiarisation and demonstration*: The task began with a brief familiarisation period where the infant was shown the trampoline and encouraged to hold it on each side. After the infant was familiar with the apparatus E1 and E2 demonstrated how a toy could be made to bounce up and down by shaking the trampoline at the rim.
- *Experimental trial one*: Following the demonstration period E1 invited the infant's participation by alternating gaze between the infant and the trampoline. If the infant was immediately successful and cooperated with the experimenter to bounce the toy on the trampoline then trial 2 was administered. However, if the infant was not successful within 30 seconds E1 and E2 carried out the demonstration phase again. Following the second demonstration the infant was encouraged to participate again

this time using verbal cues such as “Come and try” and “Look, trampoline!”. If the infant was still unsuccessful the demonstration was repeated for a third time but this time E2 also encouraged the infant to stand by her and hold the trampoline together with her. If the infant was unsuccessful after three demonstrations the task was discontinued.

- *Experimental trial two*: Following trial one, E1 briefly removed the toy then after a short period placed it back on the trampoline and invited the infant’s participation again. After five seconds of play, trial three was administered. If the infant was not successful after 60 seconds then the task was discontinued.
- *Experimental trial three*: Following trial two, E1 briefly removed the toy then after a short period placed it back on the trampoline. The infant’s participation was once again prompted by E1. However, in this trial, unlike trials one and two, when the infant picked up their side of the trampoline E1 dropped her side and placed her hands and face downwards for an interruption period of 15 seconds. Following the interruption period E1 looked back up, picked up her side, and continued as before. If the infant had disengaged, E1 prompted the infant’s participation to continue. After five seconds of play trial four was administered.
- *Experimental trial four*: The procedure for trial three was repeated.

Coding

The same coding schema used by Warneken and Tomasello (2006, 2007) was used to code infant’s behaviour. For each trial of the *Tubes-with-handles* and *Trampoline* task infant’s behaviour was coded according to their level of Coordination/Engagement and their behaviour during each interruption period (Tables 1 and 2).

Table 1. Coding schema for performance on the *Cooperation: Tubes-with-handles* task

Coordination	
Category	***Definition***
No success (Score=0)	Tubes not opened
Uncoordinated (Score=1)	Success after more than 5 seconds of inappropriate actions such as standing on wrong side, letting tube drop more than once, individual play, or individual attempts
Coordinated (Score=2)	Success, but some inappropriate actions, but not for more than 5 seconds; releasing handle not more than once
Very coordinated (Score=3)	Success after immediate understanding of their role. Infant positions herself in correct location and performs the correct action without mistakes.
Behaviour during interruption period	
Category	***Definition***
Disengagement	Infant leaves apparatus or plays without pursuing the goal by banging the apparatus, climbing on it, etc.
Individual attempt	Infant attempts to retrieve the object individually (infant attempts to hold both handles or peel it open on one side) or attempts to continue the game alone.
Waiting	Infant remains on correct side of the apparatus, ready to perform their role
Re-engagement	Infant is ready to perform their role and in addition tries to re-engage E1, e.g. pushing the tube, pointing at the object and vocalising whilst looking at the partner.

Table 2. Coding schema for the *Cooperation: Trampoline'* task

Engagement	
Category	***Definition***
No success (Score=0)	Infant does not hold and lift trampoline
Low engagement (Score=1)	Joint play but lots of stopping and not too excited. Infant needs a lot of persuasion.
Medium engagement (Score=2)	Some stopping or not too excited.
High engagement (Score=3)	Continuous play and rather excited (placing block on trampoline; initiating play; active shaking)
Behaviour during interruption period	
Category	***Definition***
Disengagement	Infant leaves apparatus or plays without pursuing the goal by banging the apparatus, climbing on it, etc.
Individual attempt	Infant attempts to retrieve the object individually (infant attempts to hold both handles or peel it open on one side) or attempts to continue the game alone.
Waiting	Infant remains on correct side of the apparatus, ready to perform their role
Re-engagement	Infant is ready to perform their role and in addition tries to re-engage E1, e.g. pushing the tube, pointing at the object and vocalising whilst looking at the partner.

Pass/fail criteria and rationale

For each infant a median Cooperation/Engagement score was calculated across the trials. In each trial 'no success' received a score of zero, 'uncoordinated'/'low engagement' received a score of one, 'coordinated'/'medium engagement' received a score of two, and 'very coordinated'/'high engagement' received a score of three. Infants were coded as having passed the *Tubes-with-handles'* task if they showed at least one re-engagement attempt during interruption periods *and* their median Cooperation score was three. Similarly, infants were coded as having passed the *Trampoline* task if they made at least one re-engagement attempt during interruption trials *and* their median Engagement score was three. These criteria were decided upon for three reasons. Firstly, as highlighted in previous literature the re-engagement attempt provided the indication that the infants possessed the social cognitive understanding to form a joint goal. Secondly, in the original experimental study it was not until 24 months with the *Tubes with handles* task, and later with the *Trampoline* task, that infants were able to coordinate their actions skilfully enough to execute a joint intention reliably towards a joint goal. A median Cooperation/Engagement score of three was decided upon as this represented 'skilful and reliable coordination' and would therefore be in line with the ages of developmental accomplishments noted in the original literature.

9.2 Scoring sheet

Order 1
Re-enactment of Intended Acts
Helping Task **(control)**
Gestures
Tubes
Trampoline
Helping (experimental)

Re-enactment of Intended Acts

Trial	
Loop & Peg ***(Left, Right, Low)***	
Beads & Cup ***(Lip, front, scrape hand over top)***	
Square & Peg ***(Right, Left, Front)***	

Helping Task Control

Order 1	
Paper Balls Exp	
Pen Exp	

Gestures Task

Order 1		**R**	**L**
Experimental Point	**R**		
Control Point	**L**		
Control Gaze	**L**		
Experimental Gaze	**R**		
Experimental Point	**L**		
Control Gaze	**R**		
Control Point	**R**		
Experimental Gaze	**L**		

Cooperation Task- Tubes

No success: Tube is not being opened

Uncoordinated: Success after more than 5 seconds of inappropriate actions such as standing on wrong side, letting tube drop more than once, individual play, or individual attempts

Coordinated: Success, but some inappropriate actions, but not for more than 5 seconds; releasing handle not more than once

Very Coordinated: Success after immediate understanding of their role. Child positions herself in correct location and performs the correct action without any mistakes.

Order 1				
Tubes with Handles (cooperation)	***No Success***	***Uncoordinated***	***Coordinated***	***Very Coordinated***
Trial 1				
Trial 2				
Trial 3				
Trial 4				

Disengagement: Child leaves apparatus or plays on apparatus without pursuing the goal by banging on the apparatus, climbing on it, etc

Individual Attempts: Child attempts to retrieve the object individually in problem solving tasks (child attempts to hold both handles or peel it open on one side) or attempts to continue the game alone.

Waiting: Child remains on correct side of the apparatus and is ready to perform their role.

Reengagement: Child is ready to perform their role and in addition tries to re-engage E1, for example, by pushing the tube, pointing at the object and vocalising while looking at the partner.

Order 1				
Tubes with Handles (Interruption periods)	***Disengagement***	***Individual Attempts***	***Waiting***	***Reengagement***
Trial 3				
Trial 4				

Cooperation Task- Trampoline

No success: Child does not hold and life trampoline

Low engagement: Joint play but lots of stopping and not too excited. Child needs a lot of persuasion.

Medium engagement: Some stopping or not too excited

High engagement: Continuous play and rather excited (placing block on trampoline; initiating play; active shaking)

Order 1				
Trampoline (cooperation)	***No Success***	***Low engagement***	***Medium engagement***	***High engagement***
Trial 1				
Trial 2				
Trial 3				
Trial 4				

Disengagement: Child leaves apparatus or plays on apparatus without pursuing the goal by banging on the apparatus, climbing on it, etc.

Individual Attempts: Child attempts to retrieve the object individually in problem solving tasks (child attempts to hold both handles or peel it open on one side) or attempts to continue the game alone.

Waiting: Child remains on correct side of the apparatus and is ready to perform their role.

Reengagement: Child is ready to perform their role and in addition tries to re-engage E1, for example, by pushing the tube, pointing at the object and vocalising while looking at the partner.

Order 1				
Trampoline (Interruption periods)	***Disengagement***	***Individual Attempts***	***Waiting***	***Reengagement***
Trial 3				
Trial 4				

Helping Task Experimental

Order 1	
Paper Balls Control	
Pen Control	

Appendix 10: Confirmation of ethical approval

Document redacted