Table F-1. Key findings in studies of antipsychotics

Author, Year Study Design Groups (Dose), N Enrollment / N Final Treatment Duration/Follow-Up Time Point Post-Treatment Risk Of Bias	Mean Age, Years ± SD Mean IQ ±SD	Outcome Measure/Baseline Scores, Mean ±SD	Outcome Measure/Post-Treatment Scores, Mean ± SD
Atomoxetine			
Handen et al., 2015[1, 2] RCT G1: Atomoxetine + Parent Training, up to 1.8 mg/kg/day, 32/ G2: Atomoxetine, up to 1.8 mg/kg/day, 32/ G3: Placebo + Parent Training, up to 1.8 mg/kg/day, 32/ G4: Placebo, up to 1.8 mg/kg/day, 32/ 10 weeks/EOT Low RoB	Age G1: 8 ± 1.9 G2: 8.6 ± 2.3 G3: 7.7 ± 1.5 G4: 8.2 ± 2.4 IQ (Stanford-Binet/Mullen Scales) G1: 83.3 ± 21.6 G2: 78.7 ± 25.9 G3: 77.9 2± 5.7 G4: 86.7 ± 23.7	SNAP-IV-Parent – ADHD G1: 2.21 ± 0.38 G2: 2.18 ± 0.44 G3: 2.22 ± 0.37 G4: 2.2 ± 0.52 SNAP-IV-Parent – Inattention G1: 2.28 ± 0.46 G2: 2.3 ± 0.43 G3: 2.23 ± 0.49 G4: 2.27 ± 0.51 SNAP-IV-Parent – Hyperactivity G1: 2.14 ± 0.49 G2: 2.07 ± 0.65 G3: 2.2 ± 0.48 G4: 2.13 ± 0.69 SNAP-IV-Parent – Oppositional Defiant Disorder G1: 1.33 ± 0.74 G2: 1.31 ± 0.65 G3: 1.37 ± 0.64 G4: 1.28 ± 0.82 SNAP-IV-Teacher – ADHD G1: 1.99 ± 0.46 G2: 2 ± 0.53 G3: 1.98 ± 0.52 G4: 1.96 ± 0.59 SNAP-IV-Teacher – Inattention G1: 2.13 ± 0.52	**10 weeks** SNAP-IV-Parent – ADHD G1: 1.23 ± 0.69 G2: 1.24 ± 0.56 G3: 1.45 ± 0.62 G4: 1.74 ± 0.86 G1 vs G4, p≤0.001; ES=0.98 G2 vs G4, p≤0.001; ES=0.80 G3 vs G4, p≤0.05; ES=0.57 SNAP-IV-Parent – Inattention G1: 1.3 ± 0.72 G2: 1.36 ± 0.61 G3: 1.45 ± 0.71 G4: 1.79 ± 0.84 G1 vs G4, p≤0.001; ES=1 G2 vs G4, p≤0.001; ES=0.84 G3 vs G4, p≤0.05; ES=0.60 SNAP-IV-Parent – Hyperactivity G1: 1.15 ± 0.74 G2: 1.12 ± 0.65 G3: 1.44 ± 0.72 G4: 1.69 ± 0.97 G1 vs G4, p≤0.01; ES=0.83 G2 vs G4, p≤0.001; ES=0.68 G3 vs G4, p≤0.05; ES=0.46

Author, Year Study Design Groups (Dose), N Enrollment / N Final Treatment Duration/Follow-Up Time Point Post-Treatment Risk Of Bias	Mean Age, Years ± SD Mean IQ ±SD	Outcome Measure/Baseline Scores, Mean ±SD	Outcome Measure/Post-Treatment Scores, Mean ± SD
		G2: 2.15 ± 0.54 G3: 2.16 ± 0.46 G4: 2.28 ± 0.49 SNAP-IV-Teacher – Hyperactivity G1: 1.86 ± 0.68 G2: 1.85 ± 0.72 G3: 1.81 ± 0.79 G4: 1.64 ± 0.9 SNAP-IV-Teacher – Oppositional Defiant Disorder G1: 1.07 ± 0.73 G2: 1.21 ± 0.62 G3: 1.03 ± 0.76 G4: 1.04 ± 0.71 ABC-Parent – Irritability G1: 17.88 ± 9.25 G2: 16 ± 9.74 G3: 18.16 ± 9.24 G4: 16.97 ± 8.36 ABC-Parent – Social Withdrawal G1: 12.09 ± 7.68 G2: 12.53 ± 9.22 G3: 8.56 ± 5.69 G4: 12.09 ± 7.12 ABC-Parent – Stereotypic Behavior G1: 5.81 ± 4.69 G2: 7.41 ± 5.5 G3: 5.34 ± 5.03 G4: 5.06 ± 5.2 ABC-Parent – Hyperactivity G1: 29.94 ± 8.74 G2: 29.34 ± 9.52 G3: 31.31 ± 8.96	SNAP-IV-Parent – Oppositional Defiant Disorder G1: 0.96 ± 0.68 G2: 0.78 ± 0.52 G3: 0.7 ± 0.55 G4: 0.79 ± 0.5 p=NR SNAP-IV-Teacher – ADHD G1: 1.14 ± 0.82 G2: 1.49 ± 0.74 G3: 1.46 ± 0.82 G4: 1.4 p=NR 4 ± 0.85 SNAP-IV-Teacher – Inattention G1: 1.3 ± 0.85 G2: 1.66 ± 0.78 G3: 1.64 ± 0.82 G4: 1.63 ± 0.98 p=NR SNAP-IV-Teacher – Hyperactivity G1: 0.98 ± 0.92 G2: 1.32 ± 0.92 G3: 1.28 ± 0.99 G4: 1.25 ± 0.92 p=NR SNAP-IV-Teacher – Oppositional Defiant Disorder G1: 0.71 ± 0.65 G2: 0.87 ± 0.77 G3: 0.56 ± 0.66 G4: 0.83 ± 0.84 p=NR ABC-Parent – Irritability

Author, Year Study Design Groups (Dose), N Enrollment / N Final Treatment Duration/Follow-Up Time Point Post-Treatment Risk Of Bias	Mean Age, Years ± SD Mean IQ ±SD	Outcome Measure/Baseline Scores, Mean ±SD	Outcome Measure/Post-Treatment Scores, Mean ± SD
		G4: 31.22 ± 10.11 ABC-Parent – Inappropriate Speech G1: 5.56 ± 3.17 G2: 5.53 ± 3.44 G3: 5.72 ± 2.7 G4: 5.69 ± 3.75 ABC-Teacher – Irritability G1: 13.03 ± 10.85 G2: 12.96 ± 8.53 G3: 12.13 ± 9.62 G4: 14.71 ± 9.86 ABC-Teacher – Social Withdrawal G1: 12.1 ± 9.92 G2: 11.23 ± 7.58 G3: 9.65 ± 8.03 G4: 11.06 ± 7.48 ABC-Teacher – Stereotypic Behavior G1: 6.27 ± 5.21 G2: 6.84 ± 5.69 G3: 5.87 ± 4.19 G4: 6.32 ± 5.41 ABC-Teacher – Hyperactivity G1: 24.37 ± 11 G2: 25.58 ± 1.08 G3: 24.29 ± 11.58 G4: 25.55 ± 12.99 ABC-Teacher – Inappropriate Speech G1: 3.97 ± 3.59 G2: 4.42 ± 3.91 G3: 3.55 ± 2.96 G4: 4.71 ± 3.43	G1: 11.71 ± 8.96 G2: 10.31 ± 7.87 G3: 9.92 ± 7.94 G4: 12.95 ± 7.43 G1 vs G4, p=NR G2 vs G4, p=NR G3 vs G4, p≤0.05; ES=0.56 ABC-Parent – Social Withdrawal G1: 7.46 ± 7.55 G2: 7.57 ± 7.8 G3: 3.64 ± 4.18 G4: 6.57 ± 4.46 p=NR ABC-Parent – Stereotypic Behavior G1: 2.21 ± 2.48 G2: 4.55 ± 5.16 G3: 2.76 ± 2.73 G4: 3.52 ± 4.45 p=NR ABC-Parent – Hyperactivity G1: 16.92 ± 11.17 G2: 15.55 ± 9.77 G3: 20.24 ± 11.16 G4: 24.24 ± 13.78 G1 vs G4, p≤0.01; ES=0.69 G2 vs G4, p≤0.01; ES=0.64 G3 vs G4, p=NR ABC-Parent – Inappropriate Speech G1: 3.58 ± 2.93 G2: 3.93 ± 3.51 G3: 3.92 ± 2.33 G4: 5.24 ± 4.15 G1 vs G4, p≤0.05;

Author, Year Study Design Groups (Dose), N Enrollment / N Final Treatment Duration/Follow-Up Time Point Post-Treatment Risk Of Bias	Mean Age, Years ± SD Mean IQ ±SD	Outcome Measure/Baseline Scores, Mean ±SD	Outcome Measure/Post-Treatment Scores, Mean ± SD
		CGI-ADHD-S-Moderate G1: 9 (28) G2: 9 (28) G3: 7 (22) G4: 9 (28) CGI-ADHD-S-Marked G1: 19 (59) G2: 15 (47) G3: 18 (58) G4: 12 (38) CGI-ADHD-S-Severe/Extreme G1: 4 (13) G2: 8 (25) G3: 6 (19) G4: 11 (34) CGI-Noncompliance-S-Moderate G1: 2 (6) G2: 1 (3) G3: 1 (3) G4: 1 (3) CGI-Noncompliance-S-Marked G1: 19 (59) G2: 16 (50) G3: 17 (53) G4: 15 (47) CGI-Noncompliance-S-Severe/Extreme G1: 11 (35) G2: 15 (47) G3: 13 (41) G4: 16 (50)	ES=0.76 G2 vs G4, p=NR G3 vs G4, p≤0.05; ES=0.86 ABC-Teacher – Irritability G1: 7.65 ± 7.74 G2: 8.97 ± 6.92 G3: 6.27 ± 7.4 G4: 12.6 ± 13.06 p=NR ABC-Teacher – Social Withdrawal G1: 7.59 ± 6.61 G2: 10.54 ± 7.63 G3: 7.93 ± 8.14 G4: 8.67 ± 10.4 p=NR ABC-Teacher – Stereotypic Behavior G1: 4.24 ± 5.84 G2: 5 ± 5.3 G3: 5.2 ± 5.48 G4: 5.4 ± 6.29 p=NR ABC-Teacher – Hyperactivity G1: 13.88 ± 11.13 G2: 20.64 ± 12.52 G3: 19 ± 12.84 G4: 20.67 ± 14.18 p=NR ABC-Teacher – Inappropriate Speech G1: 2.76 ± 2.97 G2: 3.89 ± 3.38 G3: 4.13 ± 3.38 G4: 4.13 ± 3.98 G1 vs G4, p=NR

Author, Year Study Design Groups (Dose), N Enrollment / N Final Treatment Duration/Follow-Up Time Point Post-Treatment Risk Of Bias	Mean Age, Years ± SD Mean IQ ±SD	Outcome Measure/Baseline Scores, Mean ±SD	Outcome Measure/Post-Treatment Scores, Mean ± SD
		HSQ-Severity G1: 4.01 ± 1.7 G2: 3.69 ± 1.58 G3: 3.81 ± 1.49 G4: 3.99 ± 1.82 SSQ-Severity G1: 3.67 ± 2.08 G2: 3.4 ± 1.93 G3: 3.23 ± 1.93 G4: 3.75 ± 2.29	G2 vs G4, p=NR G3 vs G4, p≤0.05; ES=0.61 CGI-ADHD-I-1 or 2 G1: 15 (48.4) G2: 15 (46.9) G3: 9 (29) G4: 6 (19.4) p=NR CGI-ADHD-I-3 G1: 8 (25.8) G2: 11 (34.4) G3: 8 (25.8) G4: 5 (16.1) p=NR CGI-ADHD-I-4 G1: 6 (19.4) G2: 5 (15.6) G3: 12 (38.7) G4: 17 (54.8) p=NR CGI-ADHD-I-5 or 6 G1: 2 (6.5) G2: 1 (3.1) G3: 2 (6.5) G4: 3 (9.7) p=NR CGI-Noncompliance-I-1 or 2 G1: 10 (32.3) G2: 14 (43.8) G3: 12 (38.7) G4: 7 (22.6) p=NR CGI-Noncompliance-I-3 G1: 13 (41.9) G2: 14 (43.8)

Author, Year Study Design Groups (Dose), N Enrollment / N Final Treatment Duration/Follow-Up Time Point Post-Treatment Risk Of Bias	Mean Age, Years ± SD Mean IQ ±SD	Outcome Measure/Baseline Scores, Mean ±SD	Outcome Measure/Post-Treatment Scores, Mean ± SD
			G3: 7 (22.6) G4: 7 (22.6) p=NR CGI-Noncompliance-I-4 G1: 7 (22.6) G2: 3 (9.4) G3: 7 (22.6) G4: 12 (38.7) p=NR CGI-Noncompliance-I-5 or 6 G1: 1 (3.2) G2: 1 (3.1) G3: 5 (16.1) G4: 5 (16.1) p=NR HSQ-Severity G1: 2.38 ± 1.97 G2: 1.8 ± 1.34 G3: 2.07 ± 1.52 G4: 3 ± 1.79 G1 vs G4, p≤0.05; ES=0.46 G2 vs G4, p≤0.01; ES=0.64 G3 vs G4, p=NR SSQ-Severity G1: 2.41 ± 2.26 G2: 3.06 ± 1.78 G3: 1.6 ± 1.43 G4: 3.04 ± 2.6 G1 vs G4, p=NR G2 vs G4, p=NR G3 vs G4, p≤0.05; ES=0.47
Handen et al., 2015[1, 2] RCT			**Open Label (20 weeks/Last**

Author, Year Study Design Groups (Dose), N Enrollment / N Final Treatment Duration/Follow-Up Time Point Post-Treatment Risk Of Bias	Mean Age, Years ± SD Mean IQ ±SD	Outcome Measure/Baseline Scores, Mean ±SD	Outcome Measure/Post-Treatment Scores, Mean ± SD
G1a: Atomoxetine + Parent Training (responders) G2a: Atomoxetine (responders) G3a: Placebo + Parent Training (responders) G3b: Atomoxetine + Parent Training (placebo + PT nonresponders) G4a: Placebo (responders) G4b: Atomoxetine (placebo nonresponders) 24 Week OLE Low RoB			**Observation Carried Forward)** SNAP-IV-Parent - ADHD G3b: 1.78 ± 0.65 G4b: 2. p=NR 18 ± 0.56 SNAP-IV-Parent – Inattention G3b: 1.75 ± 0.75 G4b: 2.27 ± 0.5 p=NR SNAP-IV-Parent – Hyperactivity G3b: 1.8 ± 0.73 G4b: 2.08 ± 0.79 p=NR SNAP-IV-Parent – Oppositional Defiance Disorder G3b: 0.97 ± 0.72 G4b: 1.02 ± 0.76 p=NR HSQ-Severity G3b: 2.96 ± 1.94 G4b: 3.73 ± 1.78 p=NR ABC-Parent-Irritability G3b: 13.84 ± 10.3 G4b: 15.64 ± 9.38 p=NR ABC-Parent-Hyperactivity G3b: 28.26 ± 11.66 G4b: 30.23 ± 12.37 p=NR

Author, Year Study Design Groups (Dose), N Enrollment / N Final Treatment Duration/Follow-Up Time Point Post-Treatment Risk Of Bias	Mean Age, Years ± SD Mean IQ ±SD	Outcome Measure/Baseline Scores, Mean ±SD	Outcome Measure/Post-Treatment Scores, Mean ± SD
			Open Label (24 weeks) SNAP-IV-Parent – Total Score G1a: 2.01 ± 0.34 G2a: 2.24 ± 0.35 G3a: 2.31 ± 0.08 G4a: 1.69 ± 0.21 p=NR HSQ-Severity G1a: 3.17 ± 1.77 G2a: 3.75 ± 1.34 G3a: 4.08 ± 1.51 G4a: 2.31 ± 1.71 p=NR
Risperidone			
Lamberti et al., 2016[3] RCT G1: Aripirazole (up to 15 mg/day), 22/19 G2: Risperidone (up tp 3 mg/day), 22/18 24 weeks/EOT Low ROB	Age G1: 8.4 ± 2.9 G2: 7.9 ± 2.3 IQ NR	CGI – Severity G1: 5.4 ± 0.5 G2: 5.5 ± 0.6 CGAS – Total Score G1: 38 ± 8.3 G2: 31.42 ± 12.4 ADHD Rating Scale – Total Score G1: 39.4 ± 2.8 G2: 37.4 ± 3.9 CPRS – Hyperactivity G1: 5.7 ± 0.7 G2: 5.4 ± 0.7 CPRS – Inattention G1: 5.4 ± 0.8 G2: 4.9 ± 0.9	CGI – Improvement G1: 3 ± 1.2 G2: 2.4 ± 0.7 G1 vs G2, p=ns CGAS – Total Score G1: 51.8 ± 9.8 G2: 42.7 ± 11.5 G1 vs G2, p=ns ADHD Rating Scale – Total Score G1: 26.7 ± 7.8 G2: 29.1 ± 3 G1 vs G2, p=ns CPRS – Hyperactivity G1: 3.4 ± 1.8 G2: 3.6 ± 1.2 G1 vs G2, p=ns CPRS – Inattention G1: 3.6 ± 1.5 G2: 3.8 ± 1 G1 vs G2, p=ns

Author, Year Study Design Groups (Dose), N Enrollment / N Final Treatment Duration/Follow-Up Time Point Post-Treatment Risk Of Bias	Mean Age, Years ± SD Mean IQ ±SD	Outcome Measure/Baseline Scores, Mean ±SD	Outcome Measure/Post-Treatment Scores, Mean ± SD
[a]Scahill 2015[4] RCT G1: Risperidone (2.5 mg/day), 57/55 G2: Placebo (NA), 27/26 21 months/2 Years post-treatment Low RoB	Age G1: 8.6 ± 2.97 G2: 9.1 ± 2.58 IQ ≥70 G1: 11 (23.9) G2: 6 (13.3) <70 G1: 35 (76.1) G2: 39 (86.7)	ABC-Irritability G1: 27.22 ± 7.28 G2: 23.44 ± 7.24 ABC-Social Withdrawal G1: 16.05 ± 8.55 G2: 18.52 ± 9.72 ABC-Stereotypic Behavior G1: 10.5 ± 4.43 G2: 8.84 ± 5.22 ABC-Hyperactivity/Noncompliance G1: 34.3 ± 7.95 G2: 28.58 ± 10.4 ABC-Inappropriate Speech G1: 5.71 ± 3.93 G2: 5.59 ± 4.03 CYBOCS G1: 15.29 ± 3.15 G2: 16 ± 3.29 CGI-Severity G1: 5.09 0.7 G2: 5.23 0.65 VABS-Communication G1: 42.81 ± 14.75 G2: 44.21 ± 17.43 VABS-Daily Living Skills G1: 38.56 ± 18.44 G2: 33.25 ± 15.2 VABS-Social Skills G1: 47.4 ± 11.71 G2: 48.83 ± 18.19 VABS-Motor Skills G1: 60.34 ± 19.86 G2: 56.88 ± 24.63	**2 yrs post-treatment** ABC-Irritability G1: 14.82 ± 8.4 G2: 17.78 ± 10.82 G1 vs G2, p=0.0147 ABC-Social Withdrawal G1: 8.43 ± 6.77 G2: 13.33 ± 8.73 G1 vs G2, p=0.0130 ABC-Stereotypic Behavior G1: 6.02 ± 4.4 G2: 6.76 ± 5.37 G1 vs G2, p=0.0866 ABC-Hyperactivity/Noncompliance G1: 17.68 ± 10.16 G2: 23.38 ± 12.06 G1 vs G2, p=0.0020 ABC-Inappropriate Speech G1: 3.86 ± 3.01 G2: 5.15 ± 4.24 G1 vs G2, p=0.0433 CYBOCS G1: 11.67 ± 4.48 G2: 13.08 ± 4.6 G1 vs G2, p=0.3 CGI-Severity G1: 4.4 ± 0.89 G2: 4.65 ± 1.09 G1 vs G2, p=0.3004 **7-38 mos follow-up** VABS-Communication G1: 43.19 ± 18.45 G2: 45.21 ± 21.09 VABS-Daily Living Skills G1: 40.56 ± 20.3 G2: 38.54 ± 21.6

Author, Year Study Design Groups (Dose), N Enrollment / N Final Treatment Duration/Follow-Up Time Point Post-Treatment Risk Of Bias	Mean Age, Years ± SD Mean IQ ±SD	Outcome Measure/Baseline Scores, Mean ±SD	Outcome Measure/Post-Treatment Scores, Mean ± SD
		VABS-Adaptive Behavior G1: 41.55 ± 15.63 G2: 38.58 ± 15.38 IQ Tests G1: 50.52 ± 23.07 G2: 46.74 ± 30.57 M-RLRS – Sensory Motor G1: 6.95 ± 4.28 G2: 4.76 ± 2.98 M-RLRS – Social Relationships G1: 6.55 ± 4.29 G2: 6.28 ± 4.32 M-RLRS – Affectual Responses G1: 6.13 ± 1.9 G2: 5.2 ± 1.94 M-RLRS – Sensory Responses G1: 20.96 ± 9.06 G2: 19.24 ± 8.8 M-RLRS – Language G1: 2.67 ± 4.31 G2:4.92 ± 4.65	VABS-Social Skills G1: 50.63 ± 15.1 G2: 44.21 ± 18.07 VABS-Motor Skills G1: 61.94 ± 15.72 G2: 63.06 ± 27.17 VABS-Adaptive Behavior G1: 40.25 ± 15.51 G2: 38.96 ± 18.02 IQ Tests G1: 48.73 ± 23.92 G2: 45.85 ± 32.22 M-RLRS – Sensory Motor G1: 4.67 ± 3.06 G2: 4.2 ± 3.42 M-RLRS – Social Relationships G1: 1.02 ± 4.23 G2: 2.4 ± 3.16 M-RLRS – Affectual Responses G1: 3.49 ± 1.88 G2: 4.04 ± 1.88 M-RLRS – Sensory Responses G1: 11.02 ± 7.1 G2: 14.8 ± 7.08 M-RLRS – Language G1: -0.15 ± 3.99 G2:1.96 ± 4.3
[a]Scahill et al., 2015[5] RCT G1: Risperidone (up to 2.5 mg/kg), 49/43 G2: Placebo (NA), 52/38 8 weeks/EOT	Age G1: 8.6 ± 2.97 G2: 9.1 ± 2.58 IQ ≥70 G1: 11 (23.9) G2: 6 (13.3)	CYBOCS G1: 15.25 ± 3.25 G2: 14.8 ± 4.27	**EOT** CYBOCS G1: 11.3 ± 3.77 G2: 14.1 ± 4.43 G1 vs G2, p=0.0005; ES=0.74

Author, Year Study Design Groups (Dose), N Enrollment / N Final Treatment Duration/Follow-Up Time Point Post-Treatment Risk Of Bias	Mean Age, Years ± SD Mean IQ ±SD	Outcome Measure/Baseline Scores, Mean ±SD	Outcome Measure/Post-Treatment Scores, Mean ± SD
Low RoB	<70 G1: 35 (76.1) G2: 39 (86.7)		
[a]Scahill et al., 2015[6] RCT G1: Risperidone (up to 2.5 mg/day), 49/49 G2: Placebo (NA), 52/52 8 weeks/EOT Low RoB	Age G1: 7.5 ± 2.80 G2: 9.1 ± 2.58 IQ ≥70 G1: 23 (46.9) G2: 6 (13.3) <70 NR	ABC-Social Withdrawal/Lethargy G1: 16.4 ± 8.2 G2: 16.1 ± 8.7	**EOT** ABC-Social Withdrawal/Lethargy G1 vs G2, p=0.05; ES=0.42
[b]Scahill et al., 2015[5] RCT G1: Risperidone (up to 3.5 mg/day), 49/47 G2: Risperidone + Parent Training (up to 3.5 mg/day + mean of 10.9 sessions), 75/70 8 weeks Low RoB	Age G1: 7.5 ± 2.80 G2: 9.1 ± 2.58 IQ ≥70 G1: 23 (46.9) G2: 6 (13.3) <70 NR	CYBOCS G1: 16.2 ± 2.47 G2: 14.8 ± 4.27	**EOT** CYBOCS G1: 11 ± 3.73 G2: 14.1 ± 4.43 G1 vs G2, p=0.0001; ES=0.88
Wink et al., 2014[7] Retrospective Cohort G1: Risperidone (2.23 mg/day), 72/72 G2: Aripiprazole (11.85 mg/day), 70/70 Varies/EOT Moderate RoB	Age (at initiation) G1: 8.41 ± 3.59 G2: 9.74 ± 3.46 IQ NR	CGI-Severity NR	**EOT** CGI-Improvement G1: 3.2 ± 1.2 G2: 2.9 ± 1.2 G1 vs G2, p=ns

Author, Year Study Design Groups (Dose), N Enrollment / N Final Treatment Duration/Follow-Up Time Point Post-Treatment Risk Of Bias	Mean Age, Years ± SD Mean IQ ±SD	Outcome Measure/Baseline Scores, Mean ±SD	Outcome Measure/Post-Treatment Scores, Mean ± SD
[b]Kent et al., 2013[8] RCT G1: Risperidone (0.125-0.175 mg/day; low dose), 30/25 G2: Risperidone (1.25-1.75 mg/day; high dose), 31/25 G3: Placebo (NA), 35/27 6 weeks/EOT Low RoB	Age G1: 10 ± 3.4 G2: 9 ± 3.1 G3: 9 ± 2.6 IQ Mental age G1 + G2 + G3: 5.5 yrs	ABC-Irritability G1: 27.1 ± 6.26 G2: 28.0 ± 7.81 G3: 28.9 ± 6.10 CGI – Severity G1: 5.1 ± 0.92 G2: 5.0 ± 0.78 G3: 4.9 ± 0.67	**EOT** Mean change in: ABC-Irritability G1: -7.4 ± 8.12 G2: -12.4 ± 6.52 G3: -3.5 ± 10.67 G1 vs G3: p=ns G2 vs G3: p<0.001 CGI – Severity G1: -0.4 ± 0.73 G2: -1.0 ± 0.78 G3: -0.3 ± 0.79 G1 vs G3: p=ns G2 vs G3: p<0.001
[c]Kent et al., 2013[9] RCT G1: Risperidone (0.125-0.175 mg/day; low dose)/Risperidone, 30/25 G2: Risperidone (1.25-1.75 mg/day; high dose)/Risperidone, 31/25 G3: Placebo/Risperidone (NA), 35/27 6 weeks/End of open label trial Low RoB	Age G1: 10 ± 3.4 G2: 9 ± 3.1 G3: 9 ± 2.6 IQ Mental age G1 + G2 + G3: 5.5 yrs	ABC-Irritability G1: 13.4 ± 3.99 G2: 14.4 ± 4.64 G3: 13.7 ± 2.66 ABC-Hyperactivity G1: 30.1 ± 11.46 G2: 33.8 ± 9.75 G3: 31.4 ± 8.60 ABC-Stereotypic Behavior G1: 9.3 ± 5.17 G2: 11.5 ± 5.06 G3: 10.5 ± 5.26 ABC-Inappropriate Speech G1: 6.6 ± 3.49 G2: 7.5 ± 2.78 G3: 5.9 ± 3.42 ABC-Social Withdrawal G1: 18.2 ± 9.71 G2: 21.4 ± 9.09 G3: 18.1 ± 10.16 CYBOCS G1: 13.4 ± 3.99 G2: 14.4 ± 4.64 G3: 13.7 ± 2.66 CGI-Severity G1: 5.1 ± 0.93	**At OLE Endpoint** Mean change score ABC-Irritability G1: -13.2 ± 9.29 G2: -13 ± 10.55 G3: -11.8 ± 7.68 ABC-Hyperactivity G1: -10.5 ± 12.42 G2: -12.3 ± 11.78 G3: -11.7 ± 8.54 ABC-Stereotypic Behavior G1: -4.2 ± 6.51 G2: -4.6 ± 5.14 G3: -2.8 ± 4.12 ABC-Inappropriate Speech G1: -1.8 ± 3.93 G2: -2.1 ± 3.07 G3: -1.5 ± 2.69 ABC-Social Withdrawal G1: -8.3 ± 9.03 G2: -10.4 ± 8.57 G3: -6.9 ± 8.08 CYBOCS G1: -4.1 ± 3.01 G2: -5.6 ± 4.66 G3: -5.3 ± 4.99

Author, Year Study Design Groups (Dose), N Enrollment / N Final Treatment Duration/Follow-Up Time Point Post-Treatment Risk Of Bias	Mean Age, Years ± SD Mean IQ ±SD	Outcome Measure/Baseline Scores, Mean ±SD	Outcome Measure/Post-Treatment Scores, Mean ± SD
		G2: 5 ± 0.75 G3: 4.9 ± 0.67	CGI-Severity G1: -1 ± 1.02 G2: -1.3 ± 1.17 G3:-0.9 ± 0.88 CGI-Much or Very Much Improved G1: 14 (58) G2: 15 (60) G3: 20 (69) *% of patients with CGI-C ratings of "much" or "very much" improved at OLE endpoint was not significant among the 3 groups (p=0.684) *The percentage of patients with ABC-I response at OLE endpoint (defined as ‡ 25% improvement from DB baseline) was not statistically significantly different among the three groups (p = 0.800). **Mean change scores from baseline** ABC-Irritability G1: -7.4 ± 8.12 G2: -12.4 ± 6.52 G3: -3.5 ± 10.67 G1 Vs G3, p=0.164, ES=0.36 G2 Vs G3, p <0.001, ES=0.94 ABC-Hyperactivity G1: -10.5 ± 12.42 G2: -12.3 ± 11.78 G3: -11.7 ± 8.54 G1 Vs G3, p=0.008 G2 Vs G3, p=0.019 ABC-Stereotypy G1: -4.2 ± 6.51

Author, Year Study Design Groups (Dose), N Enrollment / N Final Treatment Duration/Follow-Up Time Point Post-Treatment Risk Of Bias	Mean Age, Years ± SD Mean IQ ±SD	Outcome Measure/Baseline Scores, Mean ±SD	Outcome Measure/Post-Treatment Scores, Mean ± SD
			G2: -4.6 ± 5.14 G3: -2.8 ± 4.12 G1 Vs G3, p=ns G2 Vs G3, p=ns ABC-Inappropriate Speech G1: -1.8 ± 3.93 G2: -2.1 ± 3.07 G3: -1.5 ± 2.69 G1 Vs G3, p=ns G2 Vs G3, p=ns ABC-Social Withdrawal/Lethargy G1: -8.3 ± 9.03 G2: -10.4 ± 8.57 G3: -6.9 ± 8.08 G1 Vs G3, p=0.817 G2 Vs G3, p=0.004 **Change from baseline** CGI-Severity G1: -0.4 ± 0.73 G2: -1 ± 0.78 G3: -0.3 ± 0.79 G1 Vs G3, p=0.769, ES=0.08 G2 Vs G3, p<0.001, ES=1.02 **EOT** CGI-Improvement – much or very much improved G1: 5 (17) G2: 18 (63) G3: 5 (15) G1 Vs G3, p=0.985 G2 Vs G3, p<0.001 CYBOCS NR

Author, Year Study Design Groups (Dose), N Enrollment / N Final Treatment Duration/Follow-Up Time Point Post-Treatment Risk Of Bias	Mean Age, Years ± SD Mean IQ ±SD	Outcome Measure/Baseline Scores, Mean ±SD	Outcome Measure/Post-Treatment Scores, Mean ± SD
[b]Scahill et al., [6]2013 RCT G1: Risperidone (up to 3.5 mg/day), 49/47 G2: Risperidone + Parent Training (up to 3.5 mg/day + mean of 10.9 sessions), 75/70 8 weeks/EOT Low RoB	Age G1: 7.5 ± 2.80 G2: 9.1 ± 2.58 IQ ≥70 G1: 23 (46.9) G2: 6 (13.3) <70 NR	ABC-Social Withdrawal/Lethargy G1: 17.1 ± 8.5 G2: 15.2 ± 9	**EOT** ABC-Social Withdrawal/Lethargy NR
[d]Shea et al., 2004[10] RCT G1: Risperidone (1 mg/mL), 24 G2: Placebo (NA), 28 8 weeks/EOT Moderate RoB	Age G1: 7.4 ± 2.7 G2: 7.1 ± 2.1 IQ - ≤84 G1: 19 (100) G2: 18 (75) IQ - >84 G1: 0 (0) G2: 6 (25)	ABC-Irritability G1: 20.6 ± 8.1 G2: 21.6 ± 10.2 ABC-Hyperactivity G1: 29.2 ± 9.5 G2: 33.6 ± 6.8 ABC-Social Withdrawal/Lethargy G1: 14 ± 6.8 G2: 13.6 ± 8.6 ABC-Stereotypy G1: 8.4 ± 5.8 G2: 9.4 ± 5.5 ABC-Inappropriate Speech G1: 4.5 ± 3.7 G2: 4.5 ± 3.7 Nisonger-Child Behavior Rating Form – Adaptive/Social G1: 3.8 ± 2.3 G2: 3.9 ± 2 Nisonger-Child Behavior Rating Form – Compliant/Calm G1: 3.8 ± 2.7 G2: 6.2 ± 2.4 Nisonger-Child Behavior Rating Form – Conduct Problem G1: 17.2 ± 8 G2: 21.5 ± 10.7	**During Treatment (1 wk)** ABC-Irritability G1: 13.7 ± 9.1 G2: 16.9 ± 10.7 **During Treatment (2 wks)** ABC-Irritability G1: 10.3 ± 7.6 G2: 15.8 ± 10.8 **During Treatment (3 wks)** ABC-Irritability G1: 8.8 ± 7.3 G2: 14.4 ± 11.7 **During Treatment (5 wks)** ABC-Irritability G1: 8.8 ± 7.2 G2: 13.7 ± 11.1 **During Treatment (8 wks)** ABC-Irritability G1: 7.5 ± 6.4 G2: 13.8 ± 12.2 **EOT** ABC-Irritability G1: 7.2 ± 5.9 G2: 14.1 ± 11.3 G1 vs G2, p=0.002; ES = -0.7 ABC-Hyperactivity

Author, Year Study Design Groups (Dose), N Enrollment / N Final Treatment Duration/Follow-Up Time Point Post-Treatment Risk Of Bias	Mean Age, Years ± SD Mean IQ ±SD	Outcome Measure/Baseline Scores, Mean ±SD	Outcome Measure/Post-Treatment Scores, Mean ± SD
		Nisonger-Child Behavior Rating Form – Hyperactive G1: 17.7 ± 5.6 G2: 19.6 ± 5.2 Nisonger-Child Behavior Rating Form – Insecure/Anxious G1: 6.3 ± 6.7 G2: 8.7 ± 6.7 Nisonger-Child Behavior Rating Form – Overly Sensitive G1: 6.7 ± 3.4 G2: 6.6 ± 3.4 Nisonger-Child Behavior Rating Form – Self-Injury/Stereotypic G1: 4.5 ± 4.4 G2: 4.1 ± 4.4 Nisonger-Child Behavior Rating Form – Self-Isolated/Ritualistic G1: 7.3 ± 4 G2: 7.8 ± 4.2 VAS-Aggression G1: 86 ± 14.5 G2: 88.3 ± 9 VAS-Defiance/Disobedience G1: 75 ± 47 G2: 94.8 ± 19.6 VAS-Hyperactivity G1: 68.7 ± 28.6 G2: 96.3 ± 4.2 VAS-Obsessive/Repetitive G1: 86.3 ± 19.4 G2: 98 ± 0	G1: 13.3 ± 8.7 G2: 26.4 ± 12.8 G1 vs G2, p=0.001; ES = -0.8 ABC-Social Withdrawal/Lethargy G1: 4.7 ± 4.4 G2: 8.2 ± 8.9 G1 vs G2, p=0.020; ES = -0.5 ABC-Stereotypy G1: 3.9 ± 4.2 G2: 6.9 ± 6.9 G1 vs G2, p=0.053; ES = -0.4 ABC-Inappropriate Speech G1: 1.9 ± 2.2 G2: 3.1 ± 3.5 G1 vs G2, p=0.058; ES = -0.4 CGI-I - much improved or very much improved G1: 14 (58.3) G2: 6 (21.4) G1 vs G2, p=0.008 Nisonger-Child Behavior Rating Form – Adaptive/Social G1: 5.3 ± 2.4 G2: 4.3 ± 2.4 G1 vs G2, p=0.072 Nisonger-Child Behavior Rating Form – Compliant/Calm G1: 8.7 ± 3.3 G2: 6.9 ± 2.9 G1 vs G2, p=0.072 Nisonger-Child Behavior Rating Form – Conduct Problem G1: 6.5 ± 5.7 G2: 15.5 ± 11.9

Author, Year Study Design Groups (Dose), N Enrollment / N Final Treatment Duration/Follow-Up Time Point Post-Treatment Risk Of Bias	Mean Age, Years ± SD Mean IQ ±SD	Outcome Measure/Baseline Scores, Mean ±SD	Outcome Measure/Post-Treatment Scores, Mean ± SD
		VAS-Tantrums/Negative Mood G1: 80.8 ± 10.1 G2:81.2 ± 13.9	G1 vs G2, p=0.005 Nisonger-Child Behavior Rating Form – Hyperactive G1: 9.4 ± 5.4 G2: 14.9 ± 8.4 G1 vs G2, p=0.021 Nisonger-Child Behavior Rating Form – Insecure/Anxious G1: 3.2 ± 4.3 G2: 5.4 ± 4.8 G1 vs G2, p=ns Nisonger-Child Behavior Rating Form – Overly Sensitive G1: 2.8 ± 2.3 G2: 4.3 ± 3.3 G1 vs G2, p=0.029 Nisonger-Child Behavior Rating Form – Self-Injury/Stereotypic G1: 2.2 ± 3.1 G2: 2.8 ± 3.9 G1 vs G2, p=ns Nisonger-Child Behavior Rating Form – Self-Isolated/Ritualistic G1: 2.4 ± 2.5 G2: 4.5 ± 5.5 G1 vs G2, p=0.078 VAS-Aggression G1: 22.3 ± 20 G2: 63.4 ± 37.3 G1 vs G2, p=0.056 VAS-Defiance/Disobedience G1: 60 ± 21.1 G2: 65.2 ± 25 G1 vs G2, p=ns

Author, Year Study Design Groups (Dose), N Enrollment / N Final Treatment Duration/Follow-Up Time Point Post-Treatment Risk Of Bias	Mean Age, Years ± SD Mean IQ ±SD	Outcome Measure/Baseline Scores, Mean ±SD	Outcome Measure/Post-Treatment Scores, Mean ± SD
			VAS-Hyperactivity G1: 39.7 ± 19.7 G2: 80.5 ± 14.2 G1 vs G2, p=0.040 VAS-Obsessive/Repetitive G1: 70 ± 16.8 G2: 48 ± 63.4 G1 vs G2, p=ns VAS-Tantrums/Negative Mood G1: 28 ± 20.9 G2:43.4 ± 28.8 G1 vs G2, p=0.496
[a]Scahill et al., 2006[11] RCT G1: Risperidone (2.5 mg/day), 48/48 16 weeks/end of 4 months Low RoB	G1: 5-16 yrs	VABS-Communication G1: 43.58 13.48 VABS-Daily Living Skills G1: 39.42 18.06 VABS-Socialization G1: 48.67 9.78 VABS-Composite G1: 42.35 15.79	**End of 4 month open label extension** VABS-Communication G1: 45.21 18.85 d=0.11 VABS-Daily Living Skills G1: 40.38 18.11 d=0.06 VABS-Socialization G1: 5.5 12.94 d=0.14 VABS-Composite G1: 42.06 13.53 d=-0.02
[a]Scahill et al., 2005[12] RCT G1: Risperidone (2.5 mg/day), 49/49 G2: Placebo (NA), 52/42 8 weeks/EOT Low RoB	Age G1 + G2: 8.82 ± 2.69 IQ NR	Ritvo-Freeman Real Life Rating Scales – Overall Score G1: 0.94 ± 0.36 G2: 1.03 ± 0.37 VABS-Total Score G1: 33.26 ± 8.38 G2: 33.51 ± 8.29	**4wks During Treatment** Ritvo-Freeman Real Life Rating Scales – Overall Score G1: 0.54 ± 0.36 G2: 0.84 ± 0.39 **EOT (8 wks)** Ritvo-Freeman Real Life Rating Scales – Overall Score G1: 0.45 ± 0.31 G2: 0.88 ± 0.4 G1 vs G2, ES=1.08

Author, Year Study Design Groups (Dose), N Enrollment / N Final Treatment Duration/Follow-Up Time Point Post-Treatment Risk Of Bias	Mean Age, Years ± SD Mean IQ ±SD	Outcome Measure/Baseline Scores, Mean ±SD	Outcome Measure/Post-Treatment Scores, Mean ± SD
			VABS-Total Score G1: 20.34 ± 7.93 G2: 30.27 ± 8.87 G1 vs G2, ES=1.03, Group X time, p<0.001
[a]Scahill et al., 2005[13] RCT G1: Risperidone (2.5 mg/day), 63/63 G2: Placebo-Substitution (NA), NA 4 weeks during open label extension/EOT Low RoB	Age G1 + G2: 8.82 ± 2.69 IQ NR	End of initial 8 wks of med exposure CGI-I – Very much improved G1: 19 (30.2) G2: ND CGI-I – Much Improved G1: 42 (66.7) G2: ND CGI-I Minimally Improved G1: 0 (0) G2: ND CGI-I – No Change G1: 2 (3.2) G2: ND CGI-I – Worse G1: 0 (0) G2: ND CGI-I – Much Worse G1: 0 (0) G2: ND ABC-Irritability G1: 9.5 ± 6.8 G2: ND ABC-Social Withdrawal/Lethargy G1: 7.3 ± 5.4 G2: ND ABC-Stereotypy G1: 4.9 ± 4.3 G2: ND ABC-Hyperactivity	**4 wks During Treatment** CGI-I – Very much improved G1: 13 (20.6) G2: ND CGI-I – Much Improved G1: 38 (60.3) G2: ND CGI-I Minimally Improved G1: 7 (11.1) G2: ND CGI-I – No Change G1: 1 (1.6) G2: ND CGI-I – Worse G1: 20 (33.9) G2: ND CGI-I – Much Worse G1: 1 (1.6) G2: ND **16 wks EOT** CGI-I – Very much improved G1: 20 (33.9) G2: ND CGI-I – Much Improved G1: 31 (5.5) G2: ND CGI-I Minimally Improved G1: 5 (8.5)

Author, Year Study Design Groups (Dose), N Enrollment / N Final Treatment Duration/Follow-Up Time Point Post-Treatment Risk Of Bias	Mean Age, Years ± SD Mean IQ ±SD	Outcome Measure/Baseline Scores, Mean ±SD	Outcome Measure/Post-Treatment Scores, Mean ± SD
		G1: 15.1 ± 10 G2: ND ABC-Inappropriate Speech G1: 3.4 ± 3.6 G2: ND	G2: ND CGI-I – No Change G1: 1 (1.7) G2: ND CGI-I – Worse G1: 1 (1.7) G2: ND CGI-I – Much Worse G1: 1 (1.7) G2: ND **EOT (Last Observation)** ABC-Irritability G1: 11.7 ± 8 G2: ND ABC-Social Withdrawal/Lethargy G1: 6.8 ± 5.9 G2: ND ABC-Stereotypy G1: 5.8 ± 4.7 G2: ND ABC-Hyperactivity G1: 15.8 ± 10.2 G2: ND ABC-Inappropriate Speech G1: 3.4 ± 3.2 G2: ND **8 wks During Treatment** Relapse Rate G1: 10 (62.5) G2: 2 (12.5)

Author, Year Study Design Groups (Dose), N Enrollment / N Final Treatment Duration/Follow-Up Time Point Post-Treatment Risk Of Bias	Mean Age, Years ± SD Mean IQ ±SD	Outcome Measure/Baseline Scores, Mean ±SD	Outcome Measure/Post-Treatment Scores, Mean ± SD
[c]Shea et al., 2004[14] RCT G1: Risperidone (0.02 mg/kg/day), 39/39 G2: Placebo (NA), 38/38 8 weeks/EOT Moderate RoB	Age G1: 7.6 ± 2.3 G2: 7.3 ± 2.3 IQ – Normal G1: 3 (9.7) G2: 11 (31.4) IQ – Borderline G1: 6 (19.4) G2: 4 (11.4) IQ – Mild G1: 12 (38.7) G2: 8 (22.9) IQ – Moderate G1: 10 (32.3) G2: 12 (34.3)	ABC-Irritability G1: 18.9 ± 8.8 G2: 21.2 ± 9.7 ABC-Hyperactivity G1: 13.7 ± 7 G2: 14.3 ± 8.2 ABC-Social Withdrawal/Lethargy G1: 27.3 ± 9.7 G2: 30.9 ± 8.8 ABC-Stereotypy G1: 4.6 ± 3.4 G2: 4.8 ± 3.7 ABC-Inappropriate Speech G1: 7.9 ± 5 G2: 8.1 ± 5.6	**Mean Change Score** ABC-Irritability G1: -12.1 ± 5.8 G2: -6.5 ± 8.4 G1 vs G2, p≤0.001 ABC-Hyperactivity G1: -8.6 ± 5.9 G2: -5.7 ± 6.9 G1 vs G2, p≤0.01 ABC-Social Withdrawal/Lethargy G1: -14.9 ± 6.7 G2: -7.4 ± 9.7 G1 vs G2, p≤0.001 ABC-Stereotypy G1: -2.6 ± 2.6 G2: -1.6 ± 3 G1 vs G2, p≤0.05 ABC-Inappropriate Speech G1: -4.3 ± 3.8 G2: -2.4 ± 4 G1 vs G2, p≤0.05
[a]Scahill et al., 2002[15] RCT G1: Risperidone (up to 2.5 mg/kg), 49/49 G2: Placebo (NA), 52/52 8 weeks/EOT Low RoB	Age G1 + G2: 8.82 ± 2.69 IQ NR	ABC-Irritability G1: 26.2 ± 7.9 G2: 25.5 ± 6.6 ABC-Lethargy/Social Withdrawal G1: 16.4 ± 8.2 G2: 16.1 ± 8.7 ABC-Stereotypic Behavior G1: 10.6 ± 4.9 G2: 9 ± 4.4 ABC-Hyperactivity G1: 31.8 ± 9.6 G2: 32.3 ± 8.5 ABC-Inappropriate Speech G1: 4.8 ± 4.1 G2: 6.5 ± 3.6	**EOT** ABC-Irritability G1: 11.3 ± 7.4 G2: 21.9 ± 9.5 G1 vs G2, p<0.001; ES=1.2 ABC-Lethargy/Social Withdrawal G1: 8.9 ± 6.4 G2: 12 ± 8.3 G1 vs G2, p=0.03, ES=0.4 ABC-Stereotypic Behavior G1: 5.8 ± 4.6 G2: 7.3 ± 4.8 G1 vs G2, p<0.001 ES=0.8 ABC-Hyperactivity G1: 17 ± 9.7

Author, Year Study Design Groups (Dose), N Enrollment / N Final Treatment Duration/Follow-Up Time Point Post-Treatment Risk Of Bias	Mean Age, Years ± SD Mean IQ ±SD	Outcome Measure/Baseline Scores, Mean ±SD	Outcome Measure/Post-Treatment Scores, Mean ± SD
		CGI-S – Moderate G1: 9 (18) G2: 9 (18) CGI-S – Marked G1: 27 (55) G2: 28 (57) CGI-S – Severe G1: 12 (24) G2: 12 (24) CGI-S – Extreme G1: 1 (2) G2: 0 (0)	G2: 27.6 ± 10.6 G1 vs G2, p<0.001 ABC-Inappropriate Speech G1: 3 ± 3.1 G2: 5.9 ± 3.8 G1 vs G2, p=0.03. ES=0.3 CGI-I – Much Improved or very much improved + 25% reduction on ABI-I G1: 34 (69) G2: 6 (12)
Nagaraj et al., 2006[16] RCT G1: Risperidone (0.5-1.0 mg/day), 19/19 G2: Placebo (NA), 20/20 6 months/EOT Moderate RoB	Age, months G1: 57.95 ± 20.84 G2: 63 ± 20.12 IQ – Borderline IQ G1: 9 (47.7) G2: 8 (40) IQ – Mild Retardation G1: 6 (31.6) G2: 5 (25) IQ – Moderate Retardation G1: 4 (21.1) G2: 7 (35)	CARS, median G1: 39.5 (32.5-46) G2: 38.5 (31.5-43) CGAS G1: 29.79 ± 7.27 G2: 32.65 ± 7.95	**EOT** CARS, median G1: 32 (24.5-40.5) G2: 37.5 (30-42.5) G1 vs G2, p<0.001 CGAS G1: 40.94 ± 7.83 G2: 35.2 ± 9.38 G1 vs G2, p=0.035 CARS (≥25% Improvement from Baseline) G1: 12 (63.2) G2: 0 (0) CGAS (≥25% Improvement from Baseline) G1: 17 (89.5) G2: 2 (10) Global Impression of Parents – Improvement to Some Extent G1: 9 (47.4) G2: 6 (30) Global Impression of Parents – Considerably Improved G1: 9 (47.4)

Author, Year Study Design Groups (Dose), N Enrollment / N Final Treatment Duration/Follow-Up Time Point Post-Treatment Risk Of Bias	Mean Age, Years ± SD Mean IQ ±SD	Outcome Measure/Baseline Scores, Mean ±SD	Outcome Measure/Post-Treatment Scores, Mean ± SD
			G2: 0 (0) Global Impression of Parents – Worsened G1: 1 (5.3) G2: 4 (20) Global Impression of Parents – No Change G1: 0 (0) G2: 9 (45)
Miral et al., 2008[17] RCT G1: Risperidone (1.2-4.0 mg/day), 15/13 G2: Haloperidol (1.0-5.7 mg/day), 15/15 12 weeks/EOT Moderate ROB	Age G1: 10 ± 2.7 G2: 10.9 ± 2.9 IQ NR	CGI-S-Mildly Ill G1: 0 (0) G2: 1 (6.7) CGI-S-Moderately Ill G1: 3 (20) G2: 2 (13.3) CGI-S-Markedly Ill G1: 4 (26.7) G2: 7 (46.7) CGI-S-Severely Ill G1: 5 (33.3) G2: 4 (26.7) CGI-S-Very Severely Ill G1: 1 (6.7) G2: 1 (6.7) ABC-Total Score G1: 85.6 ± 27.3 G2: 67.1 ± 25.1 Ritvo-Freeman Real Life Rating Scales-Social G1: 0.62 ± 0.5 G2: 0.5 ± 0.41 Ritvo-Freeman Real Life Rating Scales-Sensory Motor G1: 0.9 ± 0.52 G2: 0.69 ± 0.47 Ritvo-Freeman Real Life Rating Scales-	CGI-I-Marked Improvement G1: 2 (15.4) G2: 0 (0) G1 vs G2, p=ns CGI-I-Moderate Improvement G1: 9 (69.2) G2: 9 (60) G1 vs G2, p=ns CGI-I-Slight Improvement G1: 2 (15.4) G2: 5 (33.3) G1 vs G2, p=ns CGI-I-No Change G1: 0 (0) G2: 1 (6.7) CGI degree of improvement G1 vs G2, p=ns ABC-Total Score G1: 36.8 ± 13.8 G2: 45.8 ± 20.2 G1 vs G2, p=0.0063 Ritvo-Freeman Real Life Rating Scales-Social G1: -0.11 ± 0.38 G2: 0.02 ± 0.57 G1 vs G2, p=ns Ritvo-Freeman Real

Author, Year Study Design Groups (Dose), N Enrollment / N Final Treatment Duration/Follow-Up Time Point Post-Treatment Risk Of Bias	Mean Age, Years ± SD Mean IQ ±SD	Outcome Measure/Baseline Scores, Mean ±SD	Outcome Measure/Post-Treatment Scores, Mean ± SD
		Affect G1: 1.09 ± 0.41 G2: 1.05 ± 0.61 Ritvo-Freeman Real Life Rating Scales-Sensory G1: 0.98 ± 0.46 G2: 0.86 0.44 Ritvo-Freeman Real Life Rating Scales-Language G1: 0.52 ± 0.37 G2: 0.15 ± 0.44	Life Rating Scales-Sensory Motor G1: 0.36 ± 0.34 G2: 0.5 ± 0.44 G1 vs G2, p=ns Ritvo-Freeman Real Life Rating Scales-Affect G1: 0.54 ± 0.34 G2: 0.64 ± 0.48 G1 vs G2, p=ns Ritvo-Freeman Real Life Rating Scales-Sensory G1: 0.51 ± 0.25 G2: 0.58 ± 0.49 G1 vs G2, p=ns Ritvo-Freeman Real Life Rating Scales-Language G1: 0.04 ± 0.25 G2: -0.05 ± 0.5 G1 vs G2, p=ns
Troost et al., 2005[18] RCT G1: Risperidone (0.5-3.5 mg/day), 12/12 G2: Placebo (ND), 12/12 8 weeks/EOT Low ROB	Age G1: 9.4 ± 3.4 G2: 8.7 ± 1.2 Average or above-average IQ G1: 6 (50) G2: 9 (75) Borderline IQ G1: 4 (33) G2: 3 (25) Mild or Moderate Retardation G1: 2 (17) G2: 0 (0)	ABC-Irritability G1: 11.1 ± 8.1 G2: 12.7 ± 7.7 ABC-Social Withdrawal/Lethargy G1: 5 ± 6 G2: 6.7 ± 6.9 ABC-Stereotypic Behavior G1: 2.3 ± 3.2 G2: 4.7 ± 4.3 ABC-Hyperactivity G1: 16.8 ± 11.5 G2: 15.8 ± 9.4 ABC-Inappropriate Speech G1: 3.2 ± 3.2 G2: 2.3 ± 1.9	CGI-Minimally Improved G1: 5 (42) G2: 3 (25) CGI-Much Improvement G1: 3 (25) G2: 6 (50) CGI-Very Much Improved G1: 4 (33) G2: 3 (25) ABC-Irritability G1: 12.6 ± 9.8 G2: 20.3 ± 10.2 G1 vs G2, p=0.043 ABC-Social Withdrawal/Lethargy G1: 2.8 ± 3.1 G2: 4.8 ± 3.5

Author, Year Study Design Groups (Dose), N Enrollment / N Final Treatment Duration/Follow- Up Time Point Post- Treatment Risk Of Bias	Mean Age, Years ± SD Mean IQ ±SD	Outcome Measure/Baseline Scores, Mean ±SD	Outcome Measure/Post- Treatment Scores, Mean ± SD
			G1 vs G2, p=ns ABC-Stereotypic Behavior G1: 3.3 ± 3.5 G2: 3.4 ± 4.6 G1 vs G2, p=0.305 ABC-Hyperactivity G1: 18 ± 11.8 G2: 20.8 ± 12.1 G1 vs G2, p=0.118 ABC-Inappropriate Speech G1: 3 ± 2.8 G2: 3 ± 2.3 G1 vs G2, p=0.303
Aripiprazole			
Findling et al., 2014[19] RCT G1: Aripiprazole (2-15 mg/day), 41/22 G2: Placebo (NA), 44/19 16 weeks/EOT Low RoB	Age G1: 10.1 ± 2.8 G2: 10.8 ± 2.8 IQ G1: NR G2: NR	ABC-I – adjusted mean score NR CGI - improvement scale NR	**EOT** Relapse rate G1: 35% G2: 52% G1 vs G2: p=ns Change scores in: ABC- I G1: 5.2 G2: 9.6 G1 vs G2: p=ns ABC- Hyperactivity G1: 5.0 G2: 10.3 G1 vs G2: p=0.041 ABC- Stereotypy G1: 0.8 G2: 2.8 G1 vs G2: p=0.018 ABC- Inappropriate speech G1: 0.6 G2: 2.1 G1 vs G2: p=0.013 ABC- Social withdrawal G1: 0

Author, Year Study Design Groups (Dose), N Enrollment / N Final Treatment Duration/Follow- Up Time Point Post- Treatment Risk Of Bias	Mean Age, Years ± SD Mean IQ ±SD	Outcome Measure/Baseline Scores, Mean ±SD	Outcome Measure/Post- Treatment Scores, Mean ± SD
			G2: 1.5 G1 vs G2: p=ns CGI- I G1: 4.2 G2: 4.8 G1 vs G2: p=ns
Ghanizadeh et al., 2014[20] RCT G1: Aripiprazole (1.25-10 mg/day), 29/29 G2: Risperidone (0.25-3 mg/day), 30/30 2 months/EOT Low RoB	Age G1: 9.6 ± 3.3 G2: 9.5 ± 4.6 IQ NR	ABC-Irritability G1: 26.2 ± 4.1 G2: 21.5 ± 7.4 ABC-Hyperactivity G1: 37.1 ± 7 G2: 36 ± 6.2 ABC-Lethargy G1: 27.5 ± 8.4 G2: 25.3 ± 8.9 ABC-Stereotypy G1: 13.6 ± 5.7 G2: 13.2 ± 4.2 ABC-Speech G1: 8.6 ± 3.1 G2: 8.9 ± 3.6	EOT ABC-Irritability G1: 14.6 ± 5.5 G2: 12.5 ± 5.4 G1 vs G2, p=ns ABC-Hyperactivity G1: 21.1 ± 9 G2: 19.1 ± 6.1 G1 vs G2, p=ns ABC-Lethargy G1: 17.3 ± 7.4 G2: 16.1 ± 6.9 G1 vs G2, p=ns ABC-Stereotypy G1: 8.2 ± 5 G2: 7.4 ± 3.9 G1 vs G2, p=ns ABC-Speech G1: 4.9 ± 2.3 G2: 5.7 ± 3.1 G1 vs G2, p=ns CGI-Improvement, n Much improved G1: 9 G2: 5 Minimally improved G1: 7 G2: 12 No change G1: 5 G2: 8 Minimally worse G1: 3

F-26

Author, Year Study Design Groups (Dose), N Enrollment / N Final Treatment Duration/Follow-Up Time Point Post-Treatment Risk Of Bias	Mean Age, Years ± SD Mean IQ ±SD	Outcome Measure/Baseline Scores, Mean ±SD	Outcome Measure/Post-Treatment Scores, Mean ± SD
			G2: 2 G1 vs G2: p=ns
[e]Marcus et al., 2012[21] RCT G1: Aripiprazole (2-15 mg/kg; flexibly dosed study), 34/34 G2: Aripiprazole (5 mg/kg), 43/43 G3: Aripiprazole (10 mg/kg), 49/49 G4: Aripiprazole (15 mg/kg), 41/41 G5: Placebo (flexibly dosed study), 39/39 G6: Placebo (fixed-dose study), 37/37 8 weeks/EOT Low RoB	Age G1: 9.5 ± 3.1 G2: 9.4 ± 3 IQ NR	Pediatric QoL-Combined Scales G1: 45 ± 2.9 G2: 41.5 ± 2.3 G3: 43.4 ± 2.1 G4: 38.3 ± 2.3 G5: 38.1 ± 2.7 G6: 42.8 ± 2.5	**EOT** Pediatric QoL-Combined Scales G1: 13.4 ± 1.9 G2: 14 ± 2.4 G3: 10.4 ± 2.2 G4: 18.7 ± 2.4 G5: 2 ± 1.8 G6:10.6 ± 2.6
[e]Marcus et al., 2011[22] RCT G1: De Novo Subjects (2-15 mg/kg), 84/55 G2: Prior Placebo (2-15 mg/kg), 69/37 G3: Prior Aripiprazole (2-15 mg/kg), 169/107 52 weeks Low RoB	Age 9.6 (mean) IQ NR	CGI-Severity G1: 4.8 ± 1 G2: 4.2 ± 1 G3: 3.9 ± 1.1 ABC-Irritability G1: 23.2 ± 8.9 G2: 21.5 ± 9.8 G3: 15 ± 9.2 ABC-Lethargy/Social Withdrawal G1: 14.6 ± 8.6 G2: 11.3 ± 9.2 G3: 10.4 ± 8.9 ABC-Stereotypic Behavior G1: 8.1 ± 5.2 G2: 9.1 ± 5.6 G3: 6.4 ± 5.5 ABC-Hyperactivity G1: 28.4 ± 10.9 G2: 25.8 ± 13.2	**EOT** CGI-Severity G1: -1 ± 0.8 G2: -0.6 ± 1.2 G3: -0.1 ± 1 ABC-Irritability G1: -8 ± 10.1 G2: -6.1 ± 1.9 G3: 0.7 ± 10.2 ABC-Lethargy/Social Withdrawal G1: -6.4 ± 7.9 G2: -4.1 ± 7.2 G3: -2.3 ± 6.4 ABC-Stereotypic Behavior G1: -2.7 ± 3.1 G2: -1.9 ± 4.1 G3: -0.5 ± 4.4 ABC-Hyperactivity G1: -12.3 ± 8.5

Author, Year Study Design Groups (Dose), N Enrollment / N Final Treatment Duration/Follow-Up Time Point Post-Treatment Risk Of Bias	Mean Age, Years ± SD Mean IQ ±SD	Outcome Measure/Baseline Scores, Mean ±SD	Outcome Measure/Post-Treatment Scores, Mean ± SD
		G3: 18.4 ± 12 ABC-Inappropriate Speech G1: 5.8 ± 3.2 G2: 5.7 ± 4.2 G3: 4.2 ± 3.6 CYBOCS G1: 12.6 ± 4.6 G2: 12.1 ± 4 G3: 10.4 ± 3.9	G2: -9.1 ± 11.5 G3: 0.6 ± 10.3 ABC-Inappropriate Speech G1: -2 ± 2.5 G2: -1.8 ± 3 G3: -0.3 ± 2.4 CYBOCS G1: -2.8 ± 3.5 G2: -2.6 ± 5.4 G3: 0.2 ± 4.1
[e]Marcus et al., 2010[23] RCT G1: Aripiprazole (2-15 mg/kg; flexibly dosed study), 46/46 G2: Aripiprazole (5 mg/kg), 52/52 G3: Aripiprazole (10 mg/kg), 59/59 G4: Aripiprazole (15 mg/kg), 53/53 G5: Placebo (flexibly dosed study) 49/49 G6: Placebo (fixed-dose study) 49/49 8 weeks/EOT Low RoB	Age Flexibly dosed study: 9.3 ± 2.9 Fixed dose study: 9.7 ± 3.1 IQ NR	ABC-Irritability G1: 29.6 ± 1 G2: 28.3 ± 1 G3: 27.6 ± 0.9 G4: 28.3 ± 1 G5: 30.8 ± 1 G6: 26.9 ± 1 ABC-Social Withdrawal/Lethargy G1: 19.9 ± 1.6 G2: 17.7 ± 1.4 G3: 16.8 ± 1.3 G4: 18.9 ± 1.4 G5: 18.1 ± 1.6 G6: 18 ± 1.5 ABC-Stereotypic Behavior G1: 11.9 ± 0.9 G2: 11.4 ± 0.8 G3: 11.6 ± 0.8 G4: 10.7 ± 0.9 G5: 10.7 ± 0.8 G6: 10.7 ± 0.8 ABC-Hyperactivity G1: 34.1 ± 1.4 G2: 33.1 ± 1.4 G3: 33.7 ± 1.3 G4: 32.2 ± 1.4 G5: 34.7 ± 1.4 G6: 31 1. ± 4 ABC-Inappropriate Speech G1: 7 ± 0.6	Change score **EOT** ABC-Irritability G1: -12.9 ± 1.4 G2: -12.4 ± 1.4 G3: -13.2 ± 1.3 G4: -14.4 ± 1 G5: -5 ± 1.4 G6: -8.4 ± 1.4 G1 vs G6, p<0.05 G2 vs G6, p<0.05 G3 vs G6, p<0.05 G4 vs G6, p<0.05 ABC-Social Withdrawal/Lethargy G1: -7.9 ± 1.2 G2: -5.8 ± 1.2 G3: -4.9 ± 1.1 G4: -7.9 ± 1.1 G5: -62 ± 1.1 G6: -5.2 ± 1.2 G5 vs G6, p=ns ABC-Stereotypic Behavior G1: -4.8 ± 0.6 G2: -4.5 ± 0.7 G3: -4.2 ± 0.6 G4: -4.5 ± 0.7 G5: -2 ± 0.6 G6: -1.8 ± 0.7 G1 vs G6, p<0.05 G2 vs G6, p<0.05 G3 vs G6, p<0.05 G4 vs G6, p<0.05

Author, Year Study Design Groups (Dose), N Enrollment / N Final Treatment Duration/Follow-Up Time Point Post-Treatment Risk Of Bias	Mean Age, Years ± SD Mean IQ ±SD	Outcome Measure/Baseline Scores, Mean ±SD	Outcome Measure/Post-Treatment Scores, Mean ± SD
		G2: 5.8 ± 0.6 G3: 6.8 ± 0.5 G4: 6.3 ± 0.5 G5: 7 ± 0.6 G6: 5.9 ± 0.6	ABC-Hyperactivity G1: 12.7 ± 1.5 G2: -14 ± 1.6 G3: -13.3 ± 1.5 G4: -16.3 ± 1.6 G5: -2.8 ± 1.5 G6: -7.7 ± 1.7 G1 vs G6, p<0.05 G2 vs G6, p<0.05 G3 vs G6, p<0.05 G4 vs G6, p<0.05 ABC-Inappropriate Speech G1: -2.5 ± 0.4 G2: -2 ± 0.5 G3: -1.8 ± 0.4 G4: -2.3 ± 0.4 G5: -0.4 ± 0.4 G6: -1.1 ± 5 G4 vs G6, p<0.05
[e]Marcus et al., 2009[24] RCT G1: Aripiprazole (2-15 mg/kg), 47/39 G2: Placebo (NA), 51/36 8 weeks/EOT Low RoB	Age G1: 9.7 ± 3.2 G2: 8.8 ± 2.6 IQ NR	ABC-Irritability G1: 29.6 ± 6.4 G2: 30.2 ± 6.5 ABC-Hyperactivity/Noncompliance G1: 34.1 G2: 34.7 ABC-Stereotypic Behavior G1: 11.9 G2: 10.7 ABC-Inappropriate Speech G1: 7 G2: 7 CYBOCS G1: 12.8 G2: 13.7	Change Score **EOT** ABC-Irritability G1: -12.9 G2: -5 G1 vs G2, p<0.001 ABC-Hyperactivity/Noncompliance G1: -12.7 G2: -2.8 G1 vs G2, p<0.001 ABC-Stereotypic Behavior G1: -4.8 G2: -2 G1 vs G2, p<0.001 ABC-Inappropriate Speech G1: -2.5 G2: -0.4 G1 vs G2, p<0.001 CYBOCS G1: -3.8 G2: -0.8

Author, Year Study Design Groups (Dose), N Enrollment / N Final Treatment Duration/Follow-Up Time Point Post-Treatment Risk Of Bias	Mean Age, Years ± SD Mean IQ ±SD	Outcome Measure/Baseline Scores, Mean ±SD	Outcome Measure/Post-Treatment Scores, Mean ± SD
			G1 vs G2, p<0.001 CGI-Severity G1: -1.2 G2: -0.4 CGI-I – Very much improved or much improved G1: 31 (67) G2: 8 (16) CGI-I – Minimally improved G1: 7 (15) G2: 10 (20) CGI-I – No change G1: 6 (13) G2: 22 (45) CGI-I – Minimally worse G1: 2 (4) G2: 5 (10) CGI-I – Much or very much worse G1: 0 (0) G2: 4 (8)
[d]Marcus et al., 2009[25] RCT G1: Aripiprazole (5 mg/kg), 53/44 G2: Aripiprazole (10 mg/kg), 59/49 G3: Aripiprazole (15 mg/kg), 54/47 G4: Placebo (NA), 52/38 8 weeks/EOT Low RoB	Age G1: 9 ± 2.8 G2: 10 ± 3.2 G3: 9.5 ± 3.1 G4: 10.2 ± 3.1 IQ NR	ABC-Irritability G1: 28.6 ± 7.6 G2: 28.2 ± 7.4 G3: 28.9 ± 6.4 G4: 28 ± 6.9 ABC-Hyperactivity/Noncompliance G1: 33.1 ± 1.4 G2: 33.7 ± 1.3 G3: 32.2 ± 1.4 G4: 31 ± 1.4 ABC-Stereotypic Behavior G1: 11.4 ± 0.8 G2: 11.6 ± 0.8 G3: 11.6 ± 0.8 G4:10.7 ± 0.8	Change score **EOT** ABC-Irritability G1: -12.4 G2: -13.2 G3: -14.4 G4: -8.4 G1 v G4: p=0.032 G2 v G4: p=0.008 G3 v G4: p=0.001 ABC-Hyperactivity/Noncompliance G1: -14 ± 1.6 G2: -13.3 ± 1.5 G3: -16.3 ± 1.6 G4: -7.7 ± 1.7 G1 v G4: p≤0.005 G2 v G4: p≤0.05

Author, Year Study Design Groups (Dose), N Enrollment / N Final Treatment Duration/Follow-Up Time Point Post-Treatment Risk Of Bias	Mean Age, Years ± SD Mean IQ ±SD	Outcome Measure/Baseline Scores, Mean ±SD	Outcome Measure/Post-Treatment Scores, Mean ± SD
		ABC-Social Withdrawal/Lethargy G1: 17.7 ± 1.4 G2: 16.8 ± 1.3 G3: 18.9 ± 1.4 G4: 18 ± 1.5 ABC-Inappropriate Speech G1: 5.8 ± 0.6 G2: 6.8 ± 0.5 G3: 6.3 ± 0.5 G4: 5.9 ± 0.6 CYBOCS G1: 13.9 ± 0.6 G2: 13.5 ± 0.5 G3: 14.1 ± 05 G4:13.7 ± 0.6 CGI-S G1: 5 ± 0.1 G2: 4.9 ± 0.1 G3: 5.1 ± 0.1 G4: 4.7 ± 0.1	G3 v G4: p≤0.001 ABC-Stereotypic Behavior G1: -4.5 ± 0.68 G2: -4.2 ± 0.63 G3: -4.5 ± 0.66 G4:-1.8 ± 0.69 G1 v G4: p≤0.005 G2 v G4: p≤0.05 G3 v G4: p≤0.005 ABC-Social Withdrawal/Lethargy G1: -5.8 ± 1.2 G2: -4.9 ± 1.1 G3: -7.9 ± 1.1 G4:-5.2 ± 1.2 ABC-Inappropriate Speech G1: -2 ± 0.5 G2: -1.8 ± 0.4 G3: -2.3 ± 0.4 G4: -1.1 ± 0.5 CYBOCS G1: -2.6 ± 0.5 G2: -2.4 ± 0.4 G3: -3.2 ± 0.5 G4: -1.7 ± 0.5 G1 v G4:p=NS G2 v G4:p=NS G3 v G4:p≤ 0.05 CGI-S G1: -0.9 ± 0.2 G2: -1 ± 0.1 G3: -1.1 ± 0.2 G4:-0.6 ± 0.2

[a]Publication in a larger RCT [4-6, 11-13, 15, 26-30]

[b]Publication in a larger RCT [8, 9]

[c]Publication in a larger RCT [10, 14]

[d]Publication in a larger RCT [21-25, 31, 32]

ABC-Aberrant Behavior Checklist; CGI-Clinical Global Impression; VABS-Vineland Adaptive Behavior Scale; CYBOCS-Children's Yale-Brown Obsessive Compulsive Scale; M-RLRS-Modified Real Life Rating Scale; EOT-End of Treatment; OLE-Open-label Extension; VAS-Visual Analog Scale; CGAS-Children's Global Assessment Scale