S2 Table. Study characteristics, papers on psychotropics except for antidepressants.

Reference	Country/setting	Study design	Exposure/ duration	Outcome/assessor *Name of the scale*	Sample size/exposed	Effect estimate, Exposed/unexposed Adjusted effect estimates where available	Cohen's d*
Assessment using psychometric instruments							
i. Assessment by health care professional							
Infant (<2 years)							
Platt 1989 [114]	USA/Collaborative Perinatal Project	Cohort	Antipsychotic/ long term (more than 2 months)	Motor skills */BSID*	308/not reported	Percentage with failures on motor assessment *Gross motor* 82.5%/69.8% *Fine motor* Not reported	-†
Peng 2013 [113]	China/ Second Xiang-ya Hospital	Cohort	Antipsychotics/ throughout pregnancy	General development/*BSID-III*	152/76	Mean (SD) *Motor scale* 101.59 (8.53)/103.68 (7.19) *Cognitive scale* 100.99 (8.17)/103.11 (7.84) *Language scale* 97.26 (6.79)/98.18 (7.18) *Adaptive scale* 99.80 (8.56)/101.24 (5.83) *Social-emotional scale* 102.54 (9.72)/104.50 (8.63)	-0.26 -0.26 -0.13 -0.20 -0.26
Johnson 2012 [73]	USA/ Emory Women's Mental Health Program and community cohort	Cohort	Antipsychotics/ any (median 32 weeks)	Motor skills */Infant Neurological International Battery*	97/21	Mean (SE) 63.86 (1.78)/70.12 (1.03)	-0.71
Mortensen 2003 [81]	Denmark/ National Health Service in North Jutland County	Cohort, registry based	Antipsychotics/ any duration Benzodiazepines/any duration	Motor skills / *Boel test*	818/63 837/82	OR (95% CI) of abnormal test result Antipsychotics 4.1 (1.3 to 13.0) Benzodiazepines 7.7 (3.0 to 19.3)	0.78 [High] 1.13 [High]

Study	Country/Setting	Design	Exposure	Outcome/Instrument	N	Results	Effect size
Oberlander 2004 [87]	Canada/ Reproductive mental health program at British Columbia Women's Hospital	Cohort	Clonazepam in combination with SSRI/mean duration 81 to 127 days compared to exposed to SSRI only	Cognition/*BSID-II*	46/18	Mean (SD) *Motor skills* 93.1 (8.6)/91.5 (9.6) *Mental development* 97.2 (4.5)/100.7 (6.4)	0.17 -0.61
Reebye 2002 [92]	Canada/ Children's and Women's Health Centre of British Columbia	Cohort	Clonazepam in combination with SSRI/any duration compared to exposed to SSRI only	Overall development/*BSID-II*	38/14	Mean (SD) *Motor skills* 102 (6.3)/106 (5.4) *Mental development* 93 (5.1)/98 (8.1)	-0.70 -0.70
Reebye 2012 [38]	Canada/ Reproductive Mental Health Program at St. Paul's Hospital and British Columbia's Women's Hospital	Cohort	Clonazepam in combination with SSRI/throughout pregnancy compared to exposed to SSRI only	Overall development/*BSID-II*	37/15	Mean (SD) *Motor skills* 93 (8.6)/91 (9.5) *Mental development* 97 (47.5)/100 (6.1)	0.22 -0.54
Viggedal 1993 [40]	Sweden/Gothen-burg cohort	Cohort	Benzodiazepines/more than 5 DDD in early pregnancy (not further specified)	General development/*Griffiths' mental development scale I* Behaviour/*Neuropsychological assessment*	41/16	Mean (SD) 101.7 (7.9)/107.6 (8.3) Percentage with hyperactive [inactive] symptoms 31% [19%]/16% [4%] Percentage with short [long] attention span 31% [19%]/8% [4%]	-0.72 0.36 [High] [0.45] [High] 0.58 [High] [0.45] [High]
Gidai 2008a [105]	Hungary/Korányi Hospital	Cohort	Alprazolam/ high doses used for suicide attempt versus unexposed siblings	Cognition/*Hungarian development test* Behaviour/*Behavioural style questionnaire*	20/9	Mean (SD) 96.7 (6.2)/97.3 (6.1) Percentage with behavioural deviation 11%/27%	-0.10 -0.43 [High]
Gidai 2008b [106]	Hungary/Korányi Hospital	Cohort	Medazepam/ high doses used for suicide attempt versus unexposed	Cognition/*Hungarian development test* Behaviour/*Behavioural style questionnaire*	19/9	Mean (SD) 100.0 (1.0)/98.5 (4.7) Percentage with behavioural deviation 22%/10%	0.43 0.33 [High]

			siblings				
Gidai 2008c [107]	Hungary/Korányi Hospital	Cohort	Chlordiazep-oxide/high doses used for suicide attempt versus unexposed sib-lings	Cognition/*Hungarian development test* Behaviour/*Behavioural style questionnaire*	52/31	Mean (SD) 101.4 (9.8)/95.0 (2.0) Percentage with behavioural deviation 19%/24%	0.83 -0.11 [High]
Timmermann 2008b [112]	Hungary/Korányi Hospital	Cohort	Meproba-mate/high doses used for suicide attempt compared to unexposed siblings	Cognition/*Hungarian development test* Behaviour/*Behavioural style questionnaire*	56/31 54/31	Mean (SD) 99.2 (8.1)/97.8 (8.8) Percentage with behavioural deviation 26%/22%	0.17 0.10 [High]
Laegreid 1992 [108]	Sweden/Gothenburg cohort	Cohort	Benzodiazepi-nes/regular use (not defined fur-ther)	Motor skills / *Touwen Neurologic Assessment Clinical neurologic assessment*	42/16	Mean (SD) *Posture and motility* 69.56 (5.03)/72.92 (2.08) Percentage with slow [stereotyped] movements 31% [13%]/4% [0%] Percentage with normal walking pattern 63%/100%	-0.96 0.71 [High] [0.48] [High] -0.98
Petik 2008a [115]	Hungary/Korányi Hospital	Cohort	Glutethimide/high doses used for suicide attempt compared to un-exposed siblings	Cognition/*Hungarian development test* Behaviour/*Behavioural style questionnaire*	26/14 22/12	Mean (SD) 100.4 (10.8)/98.8 (9.6) Percentage with behavioural deviation 33%/30%	0.16 0.07 [High]
Petik 2008b [116]	Hungary/Korányi Hospital	Cohort	Amobarbital/high doses used for suicide attempt compared to un-exposed siblings	Cognition/*Hungarian development test* Behaviour/*Behavioural style questionnaire*	21/11 13/8	Mean (SD) 96.4 (8.4)/100.0 (12.2) Percentage with behavioural deviation 40%/60%	-0.35 -0.46 [High]
Timmermann 2008a [117]	Hungary/Korányi Hospital	Cohort	Barbital, hexobar-bital, butobarbi-tal/high doses used for suicide attempt compared	Cognition/*Hungarian development test*	8/5 8/4	Mean (SD) Barbital 90.8 (18.8)/95.0 (8.7) Hexobarbital 92.5 (8.7)/96.3 (7.5)	-0.26 -0.47

				to unexposed siblings	6/3	Butobarbital 100.0 (0)/ 100.0 (0)	undefined
				Behaviour/ *Behavioural style questionnaire*		Percentage with behavioural deviation	
					5/3	Barbital 33%/100%	-2.24 [High]
					8/4	Hexobarbital 25%/25%	0.00 [High]
					6/3	Butobarbital 33%/66%	-0.71 [High]

Preschool (2-5 years)

Hurault-Delarue 2016 [69]	France/ EFEMERIS database	Cohort, registry based	Antipsychotics/ any in 2nd or 3rd trimester	Overall development/*Compulsory medical exam*	32 373/70	Percentage with delayed development Antipsychotics *Motor* 2.1%/0.4%	0.10 [High]
						Mental 2.1%/1.0%	0.06 [High]
			Anxiolytics, including hypnotics/any in 2nd or 3rd trimester		32 303/281	Anxiolytics *Motor* 1.7%/0.4%	0.08 [High]
						Mental 1.7%/1.0%	0.04 [High]
Schechter 2017 [94]	USA/ Emory Women's Mental Health Program and community cohort	Cohort	Antipsychotics/ any duration	Cognition/*DAS*	162/not reported	Pearson correlation Antipsychotics -0.09	-†
			Anxiolytics/ any duration Hypnotics/ any duration			Anxiolytics -0.19 Hypnotics -0.04	
Hartz 1975 [35]	USA/Collaborative Perinatal Project	Cohort	Meprobamate, chlordiazepoxide/any duration	Cognition/*Stanford Binet Intelligence Scale*	28 453/694	Mean Meprobamate 96.7/96.9	-†
					28 453/435	Chlordiazepoxide 96.7/96.9	

School child (6-12 years)

Platt 1989 [114]	USA/Collaborative	Cohort	Antipsychotic/	Motor skills	308/not	Not reported, as only statistically signifi-	-†

	Perinatal Project		long term (more than 2 months)	/*Paediatric neurologic assessment*	reported	cant results are reported.	

Adolescent (13-18 years)

| Mattson 2002 [79] | USA/ California TIS and Clinical Research Program | Cohort | Benzodiazepi-nes/at least one 1st trimester exposure | Cognition/ *WISC-III* | 107/25 | Mean 100.2/105.7 | -† |

ii. Assessment by parents

Infant (<2 years)

| Reebye 2002 [92] | Canada/ Children's and Women's Health Centre of British Columbia | Cohort | Clonazepam in combination with SSRI/any duration compared to exposed to SSRI only | Behaviour/*Early Infancy Temperament Questionnaire* | 38/14 | Percentage with difficult temperament 31%/0% | 0.80 [High] |

Preschool (2-5 years)

Lupattelli 2019 [109]	Norway/MoBa	Cohort	Benzodiazepi-nes/any in mid-pregnancy (week 17-28) or late pregnancy (week 29 and later)	Motor skills/*ASQ*	4 183/28	Beta (95% CI) *Gross motor* Mid -0.47 (-0.78 to -0.17) Late 0.80 (0.12 to 1.48) *Fine motor* Mid -0.30 (-0.76 to 0.15) Late 0.74 (-0.12 to 1.60)	-0.57 [High] 0.48 [High] -0.25 [High] 0.36 [High]
				Language/*ASQ*		*Communication* Mid -0.15 (-0.43 to 0.13) Late 0.06 (-0.16 to 0.27)	-0.20 [High] 0.11 [High]
				Behaviour/*CPRS-R*		*ADHD traits* Mid -0.24 (-0.63 to 0.15) Late	-0.23 [High]

Study	Country/cohort	Design	Exposure/duration	Outcome/measure	N	Results	Effect
			z-hypnotics/ any in midpregnancy (week 17-28) or late pregnancy (week 29 and later)			0.29 (0.02 to 0.57) *Gross motor* Mid 0.03 (-0.74 to 0.80) Late 0.93 (0.28 to 1.58) *Fine motor* Mid 0.32 (-0.68 to 1.32) Late 0.58 (-0.44 to 1.59) *Communication* Mid -0.27 (-0.76 to 0.23) Late 0.49 (-0.01 to 0.98) *ADHD traits* Mid 0.11 (-0.45 to 0.67) Late -0.10 (-0.49 to 0.29)	0.43 [High] 0.02 [High] 0.57 [High] 0.13 [High] 0.23 [High] -0.23 [High] 0.42 [High] 0.08 [High] -0.10 [High]
Brandlistuen 2017 [104]	Norway/MoBa	Cohort	Benzodiazepi-nes/ any duration	Behaviour/*CBCL, short-ened*	13 734/100	Beta (95% CI) Benzodiazepines *Externalising* 0.16 (-0.17 to 0.48) *Internalising* 0.26 (0.002 to 0.52)	0.10 [High] 0.20 [High]
			z-hypnotics/any		13 693/59	z-hypnotics *Externalising* -0.52 (-0.87 to -0.15) *Internalising* 0.05 (-0.24 to 0.35)	-0.37 [High] 0.04 [High]
Misri 2006 [80]	Canada/ Reproductive mental health program at British Columbia Women's Hospital	Cohort	Clonazepam in combination with SSRI/ mean duration 137 days compared to exposed to SSRI	Behaviour and emotion-ality/*CBCL*	22/9	Mean (SD) *Internalising* 59.56 (11.20)/53.31 (10.14) *Somatic complaints* 58.78 (7.76)/54.62 (5.78) *Withdrawn*	0.59 [High] 0.63 [High]

						58.44 (9.10)/54.92 (6.22)	0.47 High
						Emotionally reactive	
						61.00 (15.17)/55.08 (7.21)	0.53 High
						Anxious/depressed	
						59.11 (10.80)/56.46 (7.46)	0.30 High
Odsbu 2015 [110]	Norway/MoBa	Cohort	Benzodiazepi-nes/one period and two/three periods	Language/ *Language competence questionnaire*	51 411/31 6	OR (95%) of low language competence One period 1.0 (0.7 to 1.3)	0.00 High
					51 174/79	Two/three periods 1.3 (0.8 to 2.3)	0.14 High

School child (6-12 years)

Radojčić 2017 [111]	The Netherlands/ Generation R	Cohort	Benzodiazepi-nes/any	Behaviour/*CBCL*	4 553/79	Beta (95% CI) *ODD subscale* 0.23 (-0.30 to 0.76)	0.10 High
						Anxiety subscale 0.35 (-0.13 to 0.83)	0.16 High

iii. Assessment by teachers/others

Preschool (2-5 years)

Misri 2006 [80]	Canada/ Reproductive mental health program at British Columbia Women's Hospital	Cohort	Clonazepam in combination with SSRI/ mean duration 137 days compared to exposed to SSRI only	Behaviour and emotion-ality/teachers *CBCL*	22/9	*Internalising* 52.17 (14.16)/53.73 (7.77)	-0.14 High
						Somatic complaints 52.33 (5.72)/52.91 (4.30)	-0.12 High
						Withdrawn 55.33 (5.68)/54.82 (4.67)	0.10 High
						Emotionally reactive 57.50 (8.19)/54.64 (7.53)	0.37 High
						Anxious/depressed 59.17 (10.40)/56.64 (7.31)	0.29 High

School child (6-12 years)

Radojčić 2017 [111]	The Netherlands/ Generation R	Cohort	Benzodiazepi-nes/any duration	Behaviour/teachers *CBCL*	3 225/62	Beta (95% CI) *ODD subscale* 0.37 (-0.07 to 0.82)	0.21 High
						Anxiety subscale 0.10 (-0.25 to 0.44)	0.07 High
				Behaviour/parent and	5 287/96	*Aggressive behaviour*	

				teacher, *CBCL*		0.75 (-0.28 to 1.77)	0.15 [High]

Assessment using medical diagnosis

Figueroa 2010 [56]	USA/MarketScan claims data	Cohort, registry based	Benzodiazepines, any	ADHD/diagnosis	36 925/311	OR (95% CI) 1.82 (0.86 to 3.85)	0.33 [High]
Janecka 2018 [72]	Israel/Meuhedet health maintenance organisation	Case-control, registry based	Lithium/any	ASD	1 426/21	HR (95% CI) 4.23 (0.58 to 30.9)	0.82 [High]

Reference numbers in brackets refer to the reference list in the article.

* Effect sizes (Cohen's d) were calculated using the metaeff package (1) for Stata (2). [High] indicates that higher/more positive values mean a higher risk among the exposed. For the remaining values, a higher value means lower risk among the exposed. Traditionally, a Cohen's d with an absolute value of 0.2 is considered a small effect, 0.5 a medium effect and 0.8 or above a large effect.(3) Only results for the oldest age band are presented here, if a paper had assessed children at multiple time points using the same outcome measure.

† Not enough information to calculate Cohen's d.

ADHD: Attention Deficit Hyperactivity Disorder, BRIEF: Behaviour Rating Inventory of Executive Function, ASQ: Ages and Stages Questionnaire, BSID: Bayley Scales of Infant Development, CBCL: Child Behaviour Checklist, CPRS:R: Conners' Parent Rating Scale, revised, DAS: Differential ability scales, MoBa: Norwegian Mother and Child Cohort, ODD: Oppositional defiant disorder, RCT: Randomised controlled trial, SDQ: Strengths and Difficulties Questionnaire, TIS: Teratogen information service, WISC: Wechsler Intelligence Scale for Children.

References:

1. Kontopantelis E, Reeves D. METAEFF: Stata module to perform effect sizes calculations for meta-analyses [Internet]. 2011 [cited 2018 Jul 6]. Available from: http://EconPapers.repec.org/RePEc:boc:bocode:s457072

2. StataCorp. Stata Statistical Software: Release 15. College Station, TX: StataCorp LP; 2017.

3. Fritz CO, Morris PE, Richler JJ. Effect size estimates: Current use, calculations, and interpretation. J Exp Psychol Gen. 2012;141(1):2–18.