Long-term Changes in Observed Behaviour after Exposure to Psychiatric Drugs
A Systematic Review of Animal Studies

Pia Brandt Danborg, Anders Lykkemark Simonsen, Asbjørn Hróbjartsson, Peter C. Gøtzsche

The Nordic Cochrane Centre, Copenhagen, Denmark.

Corresponding author: Pia Brandt Danborg, MSc.
Email: pd@cochrane.dk

Conflicts of interest: None.

Citation: Danborg PB, Simonsen AL, Hróbjartsson A, Gøtzsche PC. Long-term changes in observed behaviour after exposure to psychiatric drugs: A systematic review of animal studies. Nordic Cochrane Centre 2017; 26 October.

Abstract

Many patients are taking psychiatric drugs for years despite little knowledge about their long-term harms. Here, we summarise the findings of whether psychiatric drug exposure causes long-term harms in mammals after a drug-free period.

We searched PubMed, Biosis and Embase for controlled animal studies (no behavioural priming; behavioural assessments performed after a 90 day drug-free period).

Data was extracted by two assessors for animal characteristics, study design, funding and behavioural outcomes: sleep, addiction, pain, anxiety, depression, locomotion, cognition and social (sexual) behaviour. Meta-analysis was performed when two or more studies were eligible.

We included 33 studies in mice, rats, hamsters, cats and monkeys. The quality of the studies was poor or poorly reported and heterogeneity was high. Antidepressants caused impaired sexual behaviour; benzodiazepines, antidepressants and methyl-phenidate caused statistically significant but variable effects on anxiety. Cognition was unaffected by haloperidol, olanzapine decreased learning in one test and diazepam impaired cognition on all cognitive outcomes. Antipsychotics increased vacuous chewing movements.

Animal research on long-term outcomes is sparse and dominated by poor methodology. Still, impaired sexual behaviour, cognition and locomotion were seen. Action should be taken to improve reporting and reduce bias. We suggest that psychiatric drug research follow the animals after end of the intervention to assess long-term harms.

Introduction

The prevalence of persons diagnosed with a psychiatric disorder is increasing [1], and drugs for anxiety, depression, bipolar disorder and attention deficit hyperactivity disorder (ADHD) are widely prescribed [2-3]. Sales in Denmark of selective serotonin reuptake inhibitors (SSRIs) are so high that every seventh citizen could be treated for their entire life, although we know little about their long-term effects [4-5].

The Cochrane Collaboration publishes systematic reviews of human studies but has also called for systematic reviews of preclinical studies to improve our knowledge about drugs [6]. It should be noted, however, that methodological limitations in animal studies are common and may lead to bias and wrong conclusions [7-11]. A meta-analysis of animal studies reported that non-blinded assessments exaggerated the odds ratio (OR) by 59% on average compared to blinded assessments [12].

Animal studies on antipsychotics have reported that these drugs may lose their effect due to homoeostatic compensatory mechanisms [1,13], which supposedly cause the cognitive decline during long-term use of the drugs [14-16].

Exposure to SSRIs during foetal development may lead to lower birth weight, disturbed sleep, anhedonia and, later in life, altered sexual activity [17].

As most human studies of psychiatric drugs are short-term, it could be worthwhile to summarise the long-term outcomes in animal studies [18-19]. Thousands of preclinical studies have been performed to assess behavioural effects of psychiatric drug exposure, but most have used primed animals or have assessed only short-term outcomes. Administration of psychiatric drugs to young rats has shown that altered levels of receptors persist into adulthood, long after termination of treatment [20-22]. It is therefore possible that exposure to psychiatric drugs could lead to persisting behavioural changes.

To our knowledge, preclinical studies have not been analysed systematically for persistent behavioural changes from previous exposure to psychiatric drugs.

Aim

Assessment of whether changes in observed behaviour after exposure to psychiatric drugs persist in mammals after at least a 90-day drug-free period.

Results

We identified 7,347 unique records. We excluded 6,231 records based on the title or abstract, if the studies did not contain psychiatric drugs, a follow-up period without drug, animals, or behavioural outcomes. Records were also excluded, if the animals were primed by either drug, surgery or genetic modification for certain behaviours. Full-text screening of the remaining 1,116 articles led to exclusion of 1,083 as there was no control group, no behavioural outcomes, too short follow-up or primed animals were used. Thus, we included 33 studies.

Study characteristics

See table 1. All 33 studies were placebo-controlled, but not necessarily randomised, using the same settings and procedures for intervention and control groups, and they were published between 1981 and 2014. Twenty-seven studies were in Wistar, Long-Evans or Sprague-Dawley rats [29-56]; two studies were in monkeys [57-58]; one each in hamsters [59], cats [60] and Balb/C or Swiss Webster mice [61]. Median number of animals per study was 12 for the intervention (range 5-32) and 10 for the control (range 6-76). In nine studies, the intervention was prenatal [29,31,39,50-51,53-54,56,60].

Antidepressants were used in 16 studies, antipsychotics in nine, benzodiazepines in six and ADHD drugs in two.

Table 1. summary of the 33 included studies.

Author, year, reference	Species	Strain	Gender	Treatment days	Post-treatment days	Animals per cage	Light cycle (Light on interval)	N (drug)	N (ctrl)	Drug	Drug category	Route of adm.	Test
Andreassen & Jørgensen, 1994 (52)	Rat	Sprague-Dawley	Female	84	140	2-3	8:00-20:00	10	10	haloperidol	APS	IP	Vacuous chewing movement
Andreassen et al, 2001 (30)	Rat	Sprague-Dawley	Female	112*	196			10	5	haloperidol	APS	IM	Vacuous chewing movement
Bourke et al, 2013 (54)	Rat	Sprague-Dawley	Male	28*	90	1	7:00-19:00	6	6	escitalopram	SSRI		Social behavior Open field Acoustic startle reflex Elevated plus-maze Sucrose consumption Marble burying
Cannizzaro et al, 2001 (50)	Rat	Wistar	Both	7*	90	3	8:00-20:00	32	32	diazepam	BZD	SC	Acoustic startle reflex Open field
Cannizzaro et al, 2002 (51)	Rat	Wistar	Male	7*	90	3	8:00-20:00	72	12	alprazolam diazepam zolpidem	BZD	SC	Vogel test Auditory startle reflex Forced swimming test
Cannizzaro et al, 2005 (31)	Rat	Wistar	Male	7	240	3	8:00-20:00	6	6	diazepam	BZD	SC	Acoustic startle reflex
Dallemagne & Weiss, 1982 (61)	Mouse	BALB/c SwissWebster	Both	18	159			12	6	haloperidol	APS	SC	Open field Reinforcement test
Egan et al, 1995 (32)	Rat	Sprague-Dawley	Male	252	196	2	7:00-19:00	40	13	haloperidol	APS	IM	Vacuous chewing movement
Egan et al, 1996 (33)	Rat	Sprague-Dawley	Male	210	168	2-3	7:00-19:00	60	15	haloperidol	APS	IM	Vacuous chewing movement
Frank & Heller, 1997 (34)	Rat	Long-Evans	Male	14	110			24	8	clomipramine desipramine zimelidine	TCA TCA SSRI	IP	Total sleep
Gill et al, 2012 (57)	Monkey	Rhesus	Male	365	150	1		8	8	methylphenidate	ADHD	PO	Cocaine intake
Gray et al, 2007 (35)	Rat	Sprague-Dawley	Male	29	95	2	11:00-23:00	9	9	methylphenidate	ADHD	IP	Open field Elevated plus-maze
Gunne & Haggstrom, 1983 (40)	Rat	Sprague-Dawley	Female	210	210	1	12h cycle	17	7	haloperidol	APS	IP	Vacuous chewing movement
Gunne et al, 1986 (41)	Rat	Sprague-Dawley	Both	300	150			48	16	clozapine sulpiride thiroridazine chlorpromazine fluphenazine haloperidol	APS	PO	Vacuous chewing movement
Hansen et al, 1997 (36)	Rat	Wistar	Male	14	98	3-4	6:00-18:00	27	23	LU 10-134-C (escitalopram)	SSRI	IP	Open field Social behavior Forced swimming test
Hartley et al, 1990 (45)	Rat	Sprague-Dawley Wistar	Male	14	219	1	13:30-01:30	25	24	clomipramine	TCA	SC	Spontaneous locomotion Open field

Table 1 continued..

Author, Year, reference	Species	Strain	Gender	Treatment days	Post-treatment days	Animals per cage	Light cycle (Light on interval)	N (drug)	N (ctrl)	Drug	Drug category	Route of adm.	Test
Hilakivi & Hilakivi, 1987 (37)	Rat	Wistar	Male	12	135	6	6:00-18:00	15	7	desipramine zimelidine	TCA SSRI	IP	Forced swimming test
Hilakivi et al, 1988 (42)	Rat	Wistar	Male	12	102	6	6:00-18:00	12	6	desipramine zimelidine	TCA SSRI	IP	Auditory stimulation
Klemfuss & Gillin, 1998 (59)	Hamster	Syrian	Male	14	91	1	14h cycle, age 15 weeks lights on 24h	9	9	imipramine clomipramine	TCA	SC	Running wheel
Livezey et al, 1986 (53)	Rat	Sprague-Dawley	Male	6*	180	1	12h cycle	10	10	diazepam	BZD	SC	Tail flick test Radial arm maze
Livezey et al, 1986 (60)	Cat		NA	34*	365	1	12h cycle	8	29	diazepam	BZD		Conditioning
Marczynski et al, 1988 (29)	Rat	Sprague-Dawley	Male	21*	106	1	12h cycle	9	9	diazepam	BZD	PO	Radial arm maze
Milstein et al, 2013 (38)	Rat	Long-Evans	Male	22	161	2-3	12h cycle	11	9	olanzapine	BZD	PO	Fear conditioning Elevated plus-maze
Mirmiran et al, 1981 (43)	Rat	Wistar	Male	14	309		13:30-1:30	10	9	chlorimipramine	TCA	IP	Total sleep Mouse killing Left-right alternation
Neill et al, 1990 (48)	Rat	Long-Evans	Male	14	189		13:30-1:30	12	12	clomipramine	TCA	SC	Sexual behaviour
Olivier et al, 2011 (56)	Rat	Wistar	Male	11 *	102	2	07:00-19:00	17	21	fluoxetine	SSRI	PO	Novelty-suppressed feeding Sexual behavior Elevated plus-maze
Rodriguez-Porcel et al, 2011 (55)	Rat	Long-Evans	Both	14	132	2-3				citalopram fluoxetine	SSRI	SC	Sexual behaviour
Shrestha et al, 2014 (58)	Monkey	Rhesus	Male	365	180	1		12	12	fluoxetine	SSRI	PO	Social behaviour
Smol'nikova et al, 1985 (39)	Rat		Both	20*	90			76	32	lithium oxybutyrate lithium carbonate	APS	IP	Forced swimming test
Velazquez-Moctezuma et al, 1992 (46)	Rat	Wistar	Male	14	159		10:00-22:00	14	11	clomipramine	TCA	SC	Forced swimming test
Vogel & Hagler, 1996 (47)	Rat	Sprague-Dawley	Male	14	159			34	27	iprindole	TCA	SC	Total sleep Open field Sexual behaviour
Vogel et al, 1990 (44)	Rat	Wistar	Male	14	309			8	5	clomipramine	TCA	SC	Total sleep
Vogel et al, 1996 (49)	Rat	Long-Evans	Male	14	129	1	13:30-1:30	53	27	chlorimipramine	TCA	IP	Sexual behaviour

***** = prenatal intervention. APS – antipsychotics, BZD – benzodiazepines, ADHD – drugs targeting ADHD symptoms, TCA – tricyclic antidepressants, SSRI – selective serotonin reuptake inhibitors, IM - intramuscular, IP - intraperitoneal, PO – per oral, SC - subcutaneous. Grey box means that no data were available.

For 27 studies, more than one drug was compared to the same control group (where control groups were then split accordingly), this was not the case for the remaining six studies.

Fifteen studies described that the animals were subjected to more than one test during the follow-up, whereas 15 did not and it was unclear for three studies.

The median intervention period was 14 days (range 6-365). The median follow-up period after end of intervention was 150 days (range 90-365).

Risk of Bias

The quality of the studies was generally poor or at least poorly reported. Although it is important not only to randomise to drug and placebo but also to randomise the housing and the timing of the assessments, none of the studies were described as randomised in all three aspects. Six were described as randomised to the drug or placebo (but in one study, animals were subsequently selected based on their response to treatment for long-term follow-up) [30], three in relation to housing, and six in relation to timing.

No studies provided information about blinding of caregivers, whereas the outcome assessors were described as blinded in 16 of the 33 studies. Eleven studies reported on dropouts and in seven of these, there were fewer than 20% missing animals in the analyses. Six studies mentioned adverse effects. Twenty-one studies reported all outcomes stated in the articles' methods sections, whereas 12 studies did not. In total, data was available for 142 of 192 outcomes (74%). Twenty-nine of the 50 outcomes with no data were described as nonsignificant, whereas the remainder was not mentioned.

Of the 23 studies that provided information on sponsorship, only one was industry-funded [36]; it found no significant differences for its antidepressant drug. Three studies declared no conflicts of interest [56-58], one declared conflicting interests [54], while the remaining 29 studies had no information on this.

Binary data

Binary data was only reported for one study [55]. The results showed significantly decreased intromission behaviour for rats treated with citalopram, RR 0.36 (95% confidence interval 0.13 to 0.98) and significantly decreased ejaculation behaviour for both citalopram, RR 0.18 (0.04 to 0.73) and fluoxetine, RR 0.27 (0.08 to 0.96). Mounting behaviour was decreased for citalopram and fluoxetine, but results were only statistically significant for fluoxetine (p=.0.0319).

Continuous data

A description of the tests is shown in Supporting Information S1, and which of our eight outcomes that refers to which tests is shown in S2.

For aggression, there were no studies. For sleep, there were no significant differences with antidepressants (Fig 1, three trials), neither for studies included in the meta-analysis nor the study only stating results to be insignificant [44].

For addiction, methylphenidate showed no difference in cocaine consumption and escitalopram showed no difference in sucrose consumption (Fig 2, one trial each).

Social interaction covered also sexual behaviour (Fig 3). A total of 161 animals had received antidepressants or placebo. There were 11 outcomes, six of which addressed sexual behaviour. Sexual behaviour was impaired for all six outcomes and significantly so for number of mounts and mount latency.

Two of the social behavioural outcomes (dominance and submissive displays) were assessed in 24 monkeys, whereas the remainder were assessed in rats. Fluoxetine caused submissive behaviour, standardised mean difference (SMD) 1.09 (0.22 to 1.96) and tended also to cause dominant behaviour (p = 0.16). None of the differences in rats were statistically significant.

The outcomes for depression showed diverse results. The subgroup analysis of only antidepressants led to immobility in the forced swimming test (Fig 4), SMD 1.07 (0.58 to 1.55; 3 studies).

Tests for cognition (Fig 5) were not affected by haloperidol, whereas olanzapine decreased learning in one of three tests, SMD 2.44 (1.22 to 3.66). Diazepam impaired these functions significantly in eight of nine outcomes studied.

Benzodiazepines and similar drugs, and antidepressants and methylphenidate caused several significant effects on anxiety behaviour outcomes, sometimes increasing anxiety and sometimes decreasing it (Fig 6). This was also seen in those of the studies, where the drug was given prenatally.

For the outcome locomotion, the subgroup analysis of antipsychotics increased vacuous chewing movements (Fig 7), SMD 1.09 (0.31 to 1.87; six studies) and diazepam decreased total distance travelled, SMD 1.30 (0.44 to 2.17).

Serial data

Ten studies assessed outcomes more than once during the follow-up period and we noted whether early changes persisted. Six of the ten studies showed significant differences at the first follow-up. Four showed abnormal chewing movements on haloperidol, which remained significant in three of the studies. One study showed persistent, disturbed sexual behaviour on clomipramine, and one showed persistent impaired locomotion on desipramine, but not on zimelidine. The extracted outcomes are shown in Table 2.

Meta-regression analyses

We could not perform such analyses, as the outcomes from the individual studies were not comparable.

Other analyses

Animal settings are thought to be more homogenous than human trials, but the level of heterogeneity in our results was high: The overall I^2 values exceeded 50% for the dimensions depression, cognition and anxiety. The I^2 values were above 50% for three of six sexual behaviours, two of nine locomotive behaviours, all outcomes on depression and four of 25 outcomes on anxiety. There were many differences between studies in choice of animals and strains, housing conditions, route of administration, drug category and highly varying test conditions, even when measuring the same type of behaviour.

Sensitivity analyses (blinded observers, treatment length and drug category) and subgroup analyses (rats only, prenatal/postnatal intervention and non-skewed outcomes only) did not alter our findings or explained the heterogeneity.

Table 2: overview of studies with serial data.

Author, year (reference)	Species	Gender	Drug category	Drug	Primary outcome (test)	Outcome category	Follow-up days, first and last	
							PND28	**PND140**
Andreassen & Jørgensen, 1994 (52)	Rat	Female	APS	HAL	Vacuous chewing movements	Locomotion	P<0.05	NS
							PND28	**PND196**
Andreassen et al, 2001 (30)	Rat	Female	APS	HAL	Vacuous chewing movements	Locomotion	P<0.05	P<0.05
							PND30	**PND180**
Dallemagne & Weiss, 1982 (61)	Mouse	Female	APS	HAL	Squares entered (Open field)	Locomotion	NS	NS
		Male					NS	NS
							PND21	**PND189**
Egan et al, 1995 (32)	Rat	Male	APS	HAL	Vacuous chewing movements	Locomotion	P<0.05	P<0.05
							PND21	**PND189**
Egan et al, 1996 (33)	Rat	Male	APS	HAL	Vacuous chewing movements	Locomotion	P<0.05	P<0.05
							PND60	**PND240**
Hartley et al, 1990 (45)	Rat	Male	TCA	CLO	Total sectors entered (Open field)	Locomotion	NS	NS
					Total activity (spontaneous)		NS	NS
							PND57	**PND153**
Hilakivi & Hilakivi, 1987 (37)	Rat	Male	TCA	DES	Immobility time (Forced swimming)	Depression	P<0.05	P<0.05
			SSRI	ZIM			P<0.05	NS
							PND22-35	**PND106-119**
Klemfuss & Gillin, 1998 (59)	Hamster	Male	TCA	IMI	Amplitude (Running wheel)	Locomotion	NS	NS
				CLO			NS	NS
							>90 days	**2 months later**
Mirmiran et al, 1981 (43)	Rat	Male	TCA	CHLO	Mouse killing	Social behaviour	NS	NS
					Left-right alternation	Cognition	NS	NS
							PND90	**PND210**
Neill et al, 1990 (48)	Rat	Male	TCA	CLO	Mounts	Social behaviour (sexual)	P<0.05	P<0.05
					Intromissions		P<0.05	P<0.05
					Ejaculations		P<0.05	NS
					Mount latency		P<0.05	P<0.05
					Post-ejac pause		NS	P<0.05

Serial data for outcomes that were assessed more than once during the follow-up period after last drug dose and the last assessment was performed at 90 days or later. APS – antipsychotics, TCA – tricyclic antidepressants, SSRI – selective serotonin reuptake inhibitors, HAL – haloperidol, CLO – clomipramine, DES – desipramine, ZIM – zimelidine, IMI – imipramine, CHLO – chlorimipramine, PND – postnatal day, NS – not significant.

Discussion

We investigated whether studies in mammals with a follow-up of at least 90 days after end of intervention would show persisting harms after exposure to psychiatric drugs. To our knowledge, our systematic review is the first to assess persisting harms in animals.

Strengths and limitations

All our studies were placebo-controlled. We therefore find it likely that the assessors of outcomes were also blinded, although there was little specific information on this in the reports. There was no information on the blinding of the caregivers. If the outcome assessor and the caregiver was the same person, then the studies were not effectively blinded despite it being placebo-controlled, and the caregivers may not be effectively blinded if the colours of food, syringes needed for different cages, etc. were different.
It was also not clear whether the studies were adequately randomised with respect to distribution to groups, housing and timing of assessments.

Blinding is an important corner stone of responsible study design, as caregivers (performance bias) and observers (detection bias) tend to overestimate effects when blinding is not secured [12]. Proper randomisation is similarly important [11,62-63], but randomisation to treatment groups has often not been carried out in animal research [64-65], as animals of identical strains are thought to be similar. Depending on whether animals are bought from external companies or home-bred, there can be a specific need to ensure randomisation of animals used for a specific trial. Obviously, random housing and random timing of the assessments during a working day are also important, as differences in light intensity and temperature affect animal behaviour, as well as diurnal variations [11]. There are also factors that need to be considered when obtaining animals from external companies.

Publication bias is an important issue also in animal research and may lead to overestimation of effects by 14-45% [66-67] in systematic reviews and meta-analyses, but we did not assess this in our review due to huge heterogeneity. Selective reporting of outcomes also occurred in our study, as data was available for 74% of the outcomes mentioned in the methods section.

Theoretically, animal settings are more easily controlled than human trials, but the level of heterogeneity of our results was high. There were differences in choice of animals, gender and strains, housing conditions, route of administration, drug category and highly varying test conditions, even when measuring the same type of behaviour. For example, the outcome dimension anxiety comprised 25 different outcomes measured with eight different tests.

Most studies (27 papers) were published before 2010, where careful study design considerations were not common. The ARRIVE guidelines [24] were published in 2010 to ensure minimal research waste and heightened translational value, but they have not yet been sufficiently implemented [68]. The six studies published in 2010 or later were just as methodologically poor as the ones published earlier.

Attempts to publish our review in scientific journals

We submitted our paper to four different journals that all accept systematic reviews and animal research (Journal of Risk and Safety in Medicine, PLoS Biology, PLoS One, Psychopharmacology) and later appealed the rejection by PLoS One. The reasons that the journals declined the paper were: it was out of scope; did not add anything new; too few studies; the quality of the studies was too poor; or we did not divide the studies according to the mechanism of action of the drugs.

We consider this an unfortunate example of publication bias for which we hold the editors responsible. We cannot be blamed that the animal studies were of poor quality and it definitely has merit to publish our review, as it highlights the lack of reliable research on the long-term harms of psychiatric drugs in animals. This is important knowledge also because there are very few long-term studies in humans, and because many patients are treated for decades based on short-term trials despite the fact that this might cause permanent brain damage [5,69,70].

Our review also provides helpful directions for the planning of future, much better studies of potentially great value to patients.

Interpretation

There were persistent harms for social (including sexual) behaviour; movements; cognition, memory and learning; and depression.

We found that the animals lost interest in sexual activities when previously exposed to antidepressants, in accordance with another systematic animal review that was not limited to studies with a follow-up of at least 90 days [71]. About half of patients treated with antidepressants develop sexual disturbances [72] and there are convincing reports that these may persist years after the patients have come off the drugs [73]. Antidepressants led to immobility in the forced swimming test but how this should be interpreted in a human context is less clear, although antidepressants can be depressogenic [74].

We found that olanzapine decreased learning in one of three tests and that diazepam significantly impaired memory and learning in all tests. These findings agree with results from human studies. It is suggested that for humans, antipsychotics can cause permanent brain damage and benzodiazepines can cause impairment in anterograde memory and attention span [75,76] and dementia [77].

In monkeys, methylphenidate altered the behaviour significantly to become more submissive or dominant. ADHD drugs may cause violence in humans [78].

The vacuous chewing movements displayed by the animals is clearly equivalent to the debilitating condition tardive dyskinesia seen in humans who have been exposed to antipsychotics [5,79]. It is well-known that antipsychotics can cause tardive dyskinesia [82,83], but for newer drugs, like SSRIs and drugs for ADHD symptoms, a knowledge base of long-term harms needs to be established.

The long-term harms from the use of psychiatric drugs reported here are in accordance with findings of altered receptor levels in animals long after end of intervention [80] and also with observations in humans [5].

Given the studies' many shortcomings, our results should be interpreted cautiously, especially regarding risk of bias and reproducibility of results, but they still do provide information of how psychiatric drugs could lead to persistent harms long after end of exposure for some dimensions. It should be kept in mind that these findings are most likely underestimated due to publication bias as this is a general problem with animal research [81].

Because of several shortcomings in animal research, as also seen here, Ioannidis et al 2016 [84] suggested that preclinical research is recast with better study methods and that systematic reviews are performed prospectively based on animal studies with improved study designs.

We agree that a high proportion of animal studies are of low quality as this was also predominant in our systematic review. However, to know which animal research to repeat, one needs to do a systematic investigation, as not all animal research necessarily needs repetition in a high-quality fashion. Also, there are areas of research that are of high priority to investigate, e.g. the field of psychiatry and harms from exposure to psychiatric drugs.

We performed this systematic review of animal research to heighten focus on persistent harms as an under-researched area; we do not know much about persistent harms from exposure to psychiatric drugs.

Also, animal research is the very reason for moving into clinical trials or not. Generally, clinicians do not pay much attention to animal research, but animal studies are the qualifying point for conducting a clinical trial. And since we still have animal studies as the entry point to clinical studies, we need to draw attention to the challenges of this area regarding reporting and choice of study design and methodology. To push forward information of the need for more responsible preclinical research in psychiatry and to highlight the lack of responsible research so far, systematic reviews are still needed. Many people are taking these drugs on a daily basis, have difficulties coming off them and suffer from persistent harms after ending the exposure of drugs. Thus, we cannot dismiss the available preclinical studies in this area and let them go unhandled.

Bennani [85] has suggested areas where animal models are non-predictive of clinical outcomes; one of them is psychiatry. Animal models are often primed with drugs, operation or genetic modification and primed animals in psychiatric research reduce the clinical relevance, as face validity and construct validity are not secured [86]. Also, the predictive value from primed animals is compromised. To model a human trait or behaviour in animals genetically or with a drug, we make the animal exhibit specific psychiatric symptoms that resembles behaviours that characterise various psychiatric disorders [87,88]. But the modelling of animals is also based on the wrong premise, namely that the origin of the disorder is biological [89] or can be modelled in a single entity. The learned helplessness approach to some extent breed anxious or more sensitive mammals by exposing animals to extremely stressful situations, thus supporting the new paradigm that psychiatric disorders are contextual [90] and not genetically or biologically driven. In our review, we included only non-primed animals, being aware of the limitations of animal modelling [86], thus we were only able to include 33 very heterogeneous studies covering all psychiatric drugs ever approved, despite screening about 7,500 titles.

Green [91] has presented ten recommendations to improve translational value of animal research. Here, we elaborate on the three most relevant ones:

1. Respond to patient's needs: it is of utmost importance that the research questions are of relevance to the patients. This review highlights the extreme lack of research on long-term behavioural outcomes after previous exposure to psychiatric drugs, as many people take the drugs for years as advised by their doctor or because it is difficult to come off them. To improve animal research for the sake of the patients, we need to determine whether to continue to perform animal modelling that is based on the biological approach to psychiatry or to move towards a much better documented paradigm, that psychiatric disorders are contextual and predominantly have psychosocial causes. We show that more long-term research is urgently necessary and we believe long-term follow-up after end of exposure should be mandatory for psychiatric preclinical research.

2. Replicate animal studies independently. Unfortunately, the field is flooded by studies that are very diverse. But to change animal research and improve the translational quality, we need to assess why the studies are of poor methodology or poor reporting and why this is accepted by journal editors in the first place.

3. Review animal trials systematically: Ioannidis *et al* [84] proposed to do prospective meta-analyses. However, for this to happen, all animal studies must be registered in a central database, just like it is advised for human trials. A central register would help overcome publication bias and selective reporting. But since no trial register is currently available for animal studies, this is a proposal for the future. We furthermore recommend making ethics approval processes and review processes open and transparent and to adhere to the ARRIVE guidelines, and to set up a checking system to help secure publication of each uniquely identified study.

By knowing the harms of a psychiatric drug in the preclinical phase, less money will be wasted in the developmental phase for the pharmaceutical companies, as knowledge of harms would be prominent earlier in the developmental process.

Conclusions

Despite the enormous extent of long-term use of psychiatric drugs, animal research on long-term effects of psychiatric drugs is sparse and characterised by poor methodology and poor reporting. Impairment of sexual behaviour, cognition and locomotion was seen as well as depression, and for anxiety, several outcomes were statistically significant, but in different directions.

We strongly recommend implementation of the ARRIVE guidelines [63,92] at all institutions that are part of approving or performing animal research, as this will help improve reporting and reduce bias in preclinical studies. We suggest that multicentre studies are done to obtain sufficient power and align study methods. We also suggest that, for at least psychiatric drug research, a debate is taken on the validity of animal priming and if research is performed, that animals should be followed up long after end of the intervention to study whether drug-induced effects persist.

The protocol for the systematic review

The methodological considerations and decisions to be made in case of doubt were stated in a predefined protocol (S3).

Eligibility criteria and data extraction

We included controlled studies on psychiatric drugs on healthy mammals with a post-treatment period with no drugs of at least 90 days before assessment of behaviour.

We excluded human studies, methodological papers and studies using animal models of disease or behaviour (e.g. ailments induced by operation, genetic modification or drugs). Therefore, studies with a drug challenge prior to outcome assessment were excluded,

as the behavioural outcome will be affected by the drug. We also excluded studies testing non-medical substances, e.g. cannabis, 3,4-methylenedioxy-methamphetamine ('ecstasy') or nicotine, or reporting only on non-behavioural outcomes.

Potentially eligible articles were collected in hardcopy and read in full. References of eligible articles were scanned to locate further studies. Data was recorded in Excel.

Two unblinded investigators performed data extraction independently. Extracted data was compared and consensus reached, using a third assessor if disagreement arose. If there was any doubt about whether to include available data, we asked an arbiter who had not seen the data set.

Search strategy

We developed, through an iterative process with assistance from professionals at a university library, an extensive literature search in PubMed and Embase. Searches were performed using MeSH or Map terms in PubMed and Embase, respectively, for antidepressants and antipsychotics, benzodiazepines and the eight most sold ADHD drugs as reported by the Food and Drug Administration for 2012 [23] (atomoxetine, clonidine, dexmethylphenidate, dextroamphetamine, guanfacine, lisdexamphetamine, methylphenidate and amphetamine).

The searches combined drug categories or drug names with relevant time factor terms and relevant mammalian species (see S4; the search in Biosis was supplementary and an amendment to our protocol). The most recent searches were performed in PubMed, Embase and Biosis in August 2015. We contacted authors for outcome data that were unclear or not reported in the papers published in 2000 or later.

Quantitative data synthesis

We extracted data according to published guidelines [11,24-26]: animals (species, strain, age, sex, other conditions like rearing), study design (drug intervention, group size, dose, route of administration, housing condition, treatment length and follow-up length), bias (blinding, randomisation, excluded animals, reporting of adverse effects) and outcome measures (see details in S5).

For the meta-analyses, the outcome was behavioural change, and all behavioural assessments were included. If there were at least two eligible studies for an outcome (grouped as anxiety; depression; cognition, memory and learning; locomotion; sleep; social and sexual behaviour; aggression; pain; and addiction), meta-analysis was performed using the last time point and highest dose for the drug. If there were several active groups, we split the control group accordingly to avoid double counting.

We performed sensitivity analyses with respect to blinded observers, treatment length (30 day threshold) and drug category. Exploratory analyses were performed including non-skewed outcomes only; rats only and prenatal or postnatal intervention separately.

Risk of bias was assessed according to the methods by Syrcle (modified from Cochrane methods) [11,27] and responder selection bias was added to the bias assessment (responder selection is defined as choosing or categorising animals in the study based on tests done before beginning of the study). For animal studies, the SYRCLE risk of bias tool recommended assessment of three categories of randomisation, namely randomisation of animals to intervention or control groups; random housing to prevent behavioural differences between groups introduced through differences in light intensity and temperature; and random outcome assessments to prevent influence on results from diurnal variation.

Statistical methods

We used a random effects model. For binary data, risk ratios (RR) were reported. For the meta-analyses of continuous data and ranking scale data, the standardised mean difference (SMD) was calculated.

As there were differences in the direction of the scales, we multiplied the values by -1 as appropriate. We extracted absolute outcome values. Pooling of means and standard deviations was done as advised by the Cochrane Handbook [27](S5). Missing standard deviations were calculated or estimated by standard methods, e.g. from SEs or p-values [27]. If SD values were not estimable, results were not used for meta-analysis.

Heterogeneity was assessed as I-square and further explored by subgroup and sensitivity analyses to determine the impact of explanatory factors [25,27-28].

When we had values from more than one time point after the last drug dose, we compared informally the first follow-up value with the last one.
We planned to do meta-regression analyses, if there were at least ten comparable studies for the analyses, but this was not the case.

Acknowledgements

We thank Lisbeth W. Børgesen, MSc, PhD, university librarian and Jesper Mørch, information specialist and expert in reference tools, the University of Copenhagen, for their help with searches and reference management.

We thank Judith van Luijk and Merel Ritskes-Hoitinga from the Dutch group SYRCLE (Systematic Review Centre for Laboratory Animal Research) and Gregers Wegener and Linda Marie Kai, Translational Neuropsychiatry Unit, University of Aarhus, Denmark, for relevant and useful information and inspiring discussions during our project.

We thank Olga Ahtirschi and Karen Hovhannisyan for help with the translation of an article in Russian.

References

1. Whitaker R. The anatomy of an epidemic. In: Are psychiatric medications helpful to society? PsykoVision, Copenhagen, Denmark; 2014.
2. Gøtzsche PC. Deadly Medicines and Organised Crime: How Big Pharma Has Corrupted Healthcare. 1st ed. Vol. 1. Radcliffe Medical Press Ltd; 2013. 310 p.
3. Ilyas S, Moncrieff J. Trends in prescriptions and costs of drugs for mental disorders in England, 1998-2010. Br J Psychiatry J Ment Sci. 2012 May;200(5):393–8.
4. Gøtzsche PC. Deadly medicines and organised crime: How big pharma has corrupted health care. 1st ed. London. Radcliffe Medical Press Ltd.; 2013. 310 p.
5. Gøtzsche PC. Deadly psychiatry and organised denial. 1st ed. Copenhagen, Denmark: People's Press; 2015. 376 p.
6. Ritskes-Hoitinga M, Leenaars M, Avey M, Rovers M, Scholten R. Systematic reviews of preclinical

animal studies can make significant contributions to health care and more transparent translational medicine. Cochrane Database Syst Rev. 2014;3:ED000078.
7. Kilkenny C, Parsons N, Kadyszewski E, Festing MFW, Cuthill IC, Fry D, et al. Survey of the quality of experimental design, statistical analysis and reporting of research using animals. PloS One. 2009;4(11):e7824.
8. Peters JL, Sutton AJ, Jones DR, Rushton L, Abrams KR. A systematic review of systematic reviews and meta-analyses of animal experiments with guidelines for reporting. J Environ Sci Health B. 2006;41(7):1245–58.
9. van der Worp HB, Howells DW, Sena ES, Porritt MJ, Rewell S, O'Collins V, et al. Can animal models of disease reliably inform human studies? PLoS Med. 2010 Mar;7(3):e1000245.
10. van Luijk J, Bakker B, Rovers MM, Ritskes-Hoitinga M, de Vries RBM, Leenaars M. Systematic reviews of animal studies; missing link in translational research? PloS One. 2014;9(3):e89981.
11. Hooijmans CR, Rovers MM, de Vries RBM, Leenaars M, Ritskes-Hoitinga M, Langendam MW. SYRCLE's risk of bias tool for animal studies. BMC Med Res Methodol. 2014;14:43.
12. Bello S, Krogsbøll LT, Gruber J, Zhao ZJ, Fischer D, Hróbjartsson A. Lack of blinding of outcome assessors in animal model experiments implies risk of observer bias. J Clin Epidemiol. 2014 Jun 24;
13. Samaha A-N, Seeman P, Stewart J, Rajabi H, Kapur S. "Breakthrough" dopamine supersensitivity during ongoing antipsychotic treatment leads to treatment failure over time. J Neurosci. 2007 Mar;27(11):2979–86.
14. Ho B-C, Andreasen NC, Ziebell S, Pierson R, Magnotta V. Long-term antipsychotic treatment and brain volumes: a longitudinal study of first-episode schizophrenia. Arch Gen Psychiatry. 2011 Feb;68(2):128–37.
15. Dorph-Petersen K-A, Pierri JN, Perel JM, Sun Z, Sampson AR, Lewis DA. The influence of chronic exposure to antipsychotic medications on brain size before and after tissue fixation: a comparison of haloperidol and olanzapine in macaque monkeys. Neuropsychopharmacology. 2005 Sep;30(9):1649–61.
16. Vita A, De Peri L, Deste G, Barlati S, Sacchetti E. The Effect of Antipsychotic Treatment on Cortical Gray Matter Changes in Schizophrenia: Does the Class Matter? A Meta-analysis and Meta-regression of Longitudinal Magnetic Resonance Imaging Studies. Biol Psychiatry. 2015 Sep 15;78(6):403–12.
17. Oberlander TF, Gingrich JA, Ansorge MS. Sustained neurobehavioral effects of exposure to SSRI antidepressants during development: molecular to clinical evidence. Clin Pharmacol Ther. 2009 Dec;86(6):672–7.
18. Godfrey J. Safety of therapeutic methylphenidate in adults: a systematic review of the evidence. J Psychopharmacol Oxf Engl. 2009 Mar;23(2):194–205.
19. Deshauer D, Moher D, Fergusson D, Moher E, Sampson M, Grimshaw J. Selective serotonin reuptake inhibitors for unipolar depression: a systematic review of classic long-term randomized controlled trials. CMAJ. 2008 May;178(10):1293–301.
20. Moll GH, Hause S, Rüther E, Rothenberger A, Huether G. Early methylphenidate administration to young rats causes a persistent reduction in the density of striatal dopamine transporters. J Child Adolesc Psychopharmacol. 2001;11(1):15–24.
21. Maudhuit C, Hamon M, Adrien J. Electrophysiological activity of raphe dorsalis serotoninergic neurones in a possible model of endogenous depression. Neuroreport. 1995 Mar 7;6(4):681–4.
22. Raap DK, Garcia F, Muma NA, Wolf WA, Battaglia G, van de Kar LD. Sustained desensitization of hypothalamic 5-Hydroxytryptamine1A receptors after discontinuation of fluoxetine: inhibited neuroendocrine responses to 8-hydroxy-2-(Dipropylamino)Tetralin in the absence of changes in Gi/o/z proteins. J Pharmacol Exp Ther. 1999 Feb;288(2):561–7.
23. FDA-Approved Drugs to Treat ADHD. J Psychosoc Nurs Ment Health Serv. 2012 Mar 1;50(3):11–2.
24. Kilkenny C, Browne WJ, Cuthi I, Emerson M, Altman DG. Improving bioscience research reporting: the ARRIVE guidelines for reporting animal research. Vet Clin Pathol Am Soc Vet Clin Pathol. 2012 Mar;41(1):27–31.
25. Vesterinen HM, Sena ES, Egan KJ, Hirst TC, Churolov L, Currie GL, et al. Meta-analysis of data from animal studies: a practical guide. J Neurosci Methods. 2014 Jan 15;221:92–102.
26. Hooijmans CR, Leenaars M, Ritskes-Hoitinga M. A gold standard publication checklist to improve the quality of animal studies, to fully integrate the Three Rs, and to make systematic reviews more feasible. Altern Lab Anim ATLA. 2010 May;38(2):167–82.
27. Higgins JP, Green S. Cochrane Handbook for Systematic Reviews of Interventions [Internet]. 5.1.0 ed. The Cochrane Collaboration; 2011. Available from: http://handbook.cochrane.org/
28. Carrillo M, Ricci LA, Coppersmith GA, Melloni RH Jr. The effect of increased serotonergic neurotransmission on aggression: a critical meta-analytical review of preclinical studies. Psychopharmacology (Berl). 2009 Aug;205(3):349–68.
29. Marczynski TJ, Hawkins MC, Swann PG, Krivograd AF, Patel MK, Dugich M. Perinatal upregulation of benzodiazepine receptor ontogenesis: "fearless" and more efficient goal-directed behavior of adult rat progenies. Neurotoxicol Teratol. 1988 Apr;10(2):101–11.
30. Andreassen OA, Meshul CK, Moore C, Jørgensen HA. Oral dyskinesias and morphological changes in rat striatum during long-term haloperidol administration. Psychopharmacology 2001 June;157:11-19.
31. Cannizzaro E, Martire M, Gagliano M, Plescia F, La Barbera M, Mantia G, et al. Reversal of prenatal diazepam-induced deficit in a spatial-object learning task by brief, periodic maternal separation in adult rats. Behav Brain Res. 2005 Jun 20;161(2):320–30.
32. Egan MF, Hyde TM, Kleinman JE, Wyatt RJ. Neuroleptic induced vacuous chewing movements in rodents: Incidence and effects of long term increases in haloperidol dose. Psychopharmacology 1995;117:74-81.
33. Egan MF, Hurd Y, Ferguson J, Bachus SE, Hamid EH, Hyde TM. Pharmacological and neurochemical differences between acute and tardive vacuous chewing movements induced by haloperidol. Psychopharmacology (Berl). 1996 Oct;127(4):337–45.

34. Frank MG, Heller HC. Neonatal treatments with the serotonin uptake inhibitors clomipramine and zimelidine, but not the noradrenaline uptake inhibitor desipramine, disrupt sleep patterns in adult rats. Brain Res. 1997 Sep;768:287-93.
35. Gray JD, Punsoni M, Tabori NE, Melton JT, Fanslow V, Ward MJ, et al. Methylphenidate administration to juvenile rats alters brain areas involved in cognition, motivated behaviors, appetite, and stress. J Neurosci Off J Soc Neurosci. 2007 Jul 4;27(27):7196–207.
36. Hansen HH, Sanchez C, Meier E. Neonatal administration of the selective serotonin reuptake inhibitor Lu 10-134-C increases forced swimming-induced immobility in adult rats: A putative animal model of depression? J Pharmacol Exp Ther. 1997 Dec;283(3):1333-41.
37. Hilakivi LA, Hilakivi I. Increased adult behavioral "despair" in rats neonatally exposed to desipramine or zimeldine: an animal model of depression? Pharmacol Biochem Behav. 1987 Nov;28(3):367–9.
38. Milstein JA, Elnabawi A, Vinish M, Swanson T, Enos JK, Bailey AM, Kolb B, Frost DO. Olanzapine Treatment of Adolescent Rats Causes Enduring Specific Memory Impairments and Alters Cortical Development and Function. PLoS ONE. 2013 Feb;8(2):e57308.
39. Smol'nikova NM, Allakhverdiev VD, Liubimov BI. [Development of progeny after antenatal exposure to lithium oxybutyrate]. Farmakol Toksikol. 1985 Feb;48(1):73–6.
40. Gunne LM, Häggström JE. Reduction of nigral glutamic acid decarboxylase in rats with neuroleptic-induced oral dyskinesia. Psychopharmacology (Berl). 1983;81(3):191–4.
41. Gunne LM, Andersson U, Bondesson U, Johansson P. Spontaneous chewing movements in rats during acute and chronic antipsychotic drug administration. Pharmacol Biochem Behav. 1986 Oct;25(4):897–901.
42. Hilakivi LA, Taira T, Hilakivi I. Early postnatal deprivation of active sleep with desipramine or zimeldine impairs later behavioural reactivity to auditory stimuli in rats. Acta Physiol Scand. 1988 Feb;132(2):191–8.
43. Mirmiran M, van de Poll NE, Corner MA, van Oyen HG, Bour HL. Suppression of active sleep by chronic treatment with chlorimipramine during early postnatal development: effects upon adult sleep and behavior in the rat. Brain Res. 1981 Jan;204(1):129–46.
44. Vogel G, Neill D, Kors D, Hagler M. REM sleep abnormalities in a new animal model of endogenous depression. Neurosci Biobehav Rev. 1990;14(1):77–83.
45. Hartley P, Neill D, Hagler M, Kors D, Vogel G. Procedure- and age-dependent hyperactivity in a new animal model of endogenous depression. Neurosci Biobehav Rev. 1990;14(1):69–72.
46. Velazquez-Moctezuma J, Diaz Ruiz O. Neonatal treatment with clomipramine increased immobility in the forced swim test: an attribute of animal models of depression. Pharmacol Biochem Behav. 1992 Aug;42(4):737–9.
47. Vogel G, Hagler M. Effects of neonatally administered iprindole on adult behaviors of rats. Pharmacol Biochem Behav. 1996 Sep;55(1):157–61.
48. Neill D, Vogel G, Hagler M, Kors D, Hennessey A. Diminished sexual activity in a new animal model of endogenous depression. Neurosci Biobehav Rev. 1990 Jan;14(1):73–6.
49. Vogel G, Hagler M, Hennessey A, Richard C. Dose-dependent decrements in adult male rat sexual behavior after neonatal clorimipramine treatment. Pharmacol Biochem Behav. 1996 Jul;54(3):605–9.
50. Cannizzaro C, Martire M, Cannizzaro E, Provenzano G, Gagliano M, Carollo A, et al. Long-lasting handling affects behavioural reactivity in adult rats of both sexes prenatally exposed to diazepam. Brain Res. 2001 Jun 22;904(2):225–33.
51. Cannizzaro C, Martire M, Steardo L, Cannizzaro E, Gagliano M, Mineo A, et al. Prenatal exposure to diazepam and alprazolam, but not to zolpidem, affects behavioural stress reactivity in handling-naïve and handling-habituated adult male rat progeny. Brain Res. 2002 Oct 25;953(1-2):170–80.
52. Andreassen OA, Jørgensen HA. GM1 ganglioside attenuates the development of vacuous chewing movements induced by long-term haloperidol treatment of rats. Psychopharmacology (Berl). 1994;116:517-22.
53. Livezey GT, Marczynski TJ, Isaac L. Prenatal diazepam: chronic anxiety and deficits in brain receptors in mature rat progeny. Neurobehav Toxicol Teratol. 1986 Oct;8(5):425–32.
54. Bourke CH, Stowe ZN, Neigh GN, Olson DE, Owens MJ. Prenatal exposure to escitalopram and/or stress in rats produces limited effects on endocrine, behavioral, or gene expression measures in adult male rats. Neurotoxicol Teratol. 2013 Oct;39:100–9.
55. Rodriguez-Porcel F, Green D, Khatri N, Harris SS, May WL, Lin RCS, et al. Neonatal exposure of rats to antidepressants affects behavioral reactions to novelty and social interactions in a manner analogous to autistic spectrum disorders. Anat Rec Hoboken NJ 2007. 2011 Oct;294(10):1726–35.
56. Olivier JDA, Vallès A, van Heesch F, Afrasiab-Middelman A, Roelofs JJPM, Jonkers M, et al. Fluoxetine administration to pregnant rats increases anxiety-related behavior in the offspring. Psychopharmacology (Berl). 2011 Oct;217(3):419–32.
57. Gill KE, Pierre PJ, Daunais J, Bennett AJ, Martelle S, Gage HD, Swanson JM, Nader MA, Porrino LJ. Chronic treatment with extended release methylphenidate does not alter dopamine systems or increase vulnerability for cocaine self-administration: A study in nonhuman primates. Neuropsychopharmacology. 2012 Nov;37:2555-65.
58. Shrestha SS, Nelson EE, Liow J-S, Gladding R, Lyoo CH, Noble PL, et al. Fluoxetine administered to juvenile monkeys: effects on the serotonin transporter and behavior. Am J Psychiatry. 2014 Mar;171(3):323–31.
59. Klemfuss H, Gillin JC. Neonatal scopolamine or antidepressant treatment: Effect on development of hamster circadian rhythms. Pharmacol Biochem Behav. 1998 Feb;59(2):369-73.
60. Livezey GT, Marczynski TJ, Isaac L. Enduring effects of prenatal diazepam on the behavior, EEG, and brain receptors of the adult cat progeny. Neurotoxicology. 1986;7(2):319–33.
61. Dallemagne G, Weiss B. Altered adult behavior of mice following postnatal treatment with haloperidol. Pharmacol Biochem Behav. 1982 May;16(5):761–7.
62. Savović J, Jones HE, Altman DG, Harris RJ, Jüni P, Pildal J, et al. Influence of reported study

design characteristics on intervention effect estimates from randomized, controlled trials. Ann Intern Med. 2012 Sep 18;157(6):429–38.
63. Kilkenny C, Browne WJ, Cuthi I, Emerson M, Altman DG. Improving bioscience research reporting: the ARRIVE guidelines for reporting animal research. Vet Clin Pathol Am Soc Vet Clin Pathol. 2012 Mar;41(1):27–31.
64. Hirst JA, Howick J, Aronson JK, Roberts N, Perera R, Koshiaris C, et al. The need for randomization in animal trials: an overview of systematic reviews. PloS One. 2014;9(6):e98856.
65. Macleod MR, Fisher M, O'Collins V, Sena ES, Dirnagl U, Bath PMW, et al. Good laboratory practice: preventing introduction of bias at the bench. Stroke J Cereb Circ. 2009 Mar;40(3):e50–2.
66. Sena ES, van der Worp HB, Bath PMW, Howells DW, Macleod MR. Publication bias in reports of animal stroke studies leads to major overstatement of efficacy. PLoS Biol. 2010 Mar;8(3):e1000344.
67. Henderson VC, Demko N, Hakala A, MacKinnon N, Federico CA, Fergusson D, et al. A meta-analysis of threats to valid clinical inference in preclinical research of sunitinib. eLife. 2015;4:e08351.
68. Baker D, Lidster K, Sottomayor A, Amor S. Two years later: journals are not yet enforcing the ARRIVE guidelines on reporting standards for pre-clinical animal studies. PLoS Biol. 2014 Jan;12(1):e1001756.
69. Breggin PR. Psychiatric drug withdrawal: A guide for prescribers, therapists, patients and their families. 1st ed. New York: Springer Publishing Company; 2012. 352p.
70. Breggin PR. Brain-disabling treatments in psychiatry: Drugs, electroshock, and the psychopharmaceutical complex. 1st ed. New York: Springer Publishing Company; 2008. 576p.
71. Simonsen A, Danborg P, Gøtzsche P. Persistent sexual dysfunction after early exposure to SSRIs: systematic review of animal studies. Int J Risk Saf Med. (Accepted for publication.):2016.
72. Montejo AL, Llorca G, Izquierdo JA, Rico-Villademoros F. Incidence of sexual dysfunction associated with antidepressant agents: a prospective multicenter study of 1022 outpatients. Spanish Working Group for the Study of Psychotropic-Related Sexual Dysfunction. J Clin Psychiatry. 2001;62 Suppl 3:10–21.
73. Hogan C, Le Noury J, Healy D, Mangin D. One hundred and twenty cases of enduring sexual dysfunction following treatment. Int J Risk Saf Med. 2014;26(2):109–16.
74. Fava GA. Can long-term treatment with antidepressant drugs worsen the course of depression? J Clin Psychiatry. 2003 Feb;64(2):123–33.
75. Barbee JG. Memory, benzodiazepines, and anxiety: integration of theoretical and clinical perspectives. J Clin Psychiatry. 1993 Oct;54 Suppl:86–97; discussion 98–101.
76. Buffett-Jerrott SE, Stewart SH. Cognitive and sedative effects of benzodiazepine use. Curr Pharm Des. 2002;8(1):45–58.
77. Billioti de Gage S, Moride Y, Ducruet T, Kurth T, Verdoux H, Tournier M, et al. Benzodiazepine use and risk of Alzheimer's disease: case-control study. BMJ. 2014 Sep 9;349(sep09 2):g5205–g5205.
78. Moore TJ, Glenmullen J, Furberg CD. Prescription drugs associated with reports of violence towards others. PloS One. 2010;5(12):e15337.
79. Turrone P, Remington G, Nobrega JN. The vacuous chewing movement (VCM) model of tardive dyskinesia revisited: is there a relationship to dopamine D(2) receptor occupancy? Neurosci Biobehav Rev. 2002 May;26(3):361–80.
80. Borue X, Chen J, Condron BG. Developmental effects of SSRIs: lessons learned from animal studies. Int J Dev Neurosci Off J Int Soc Dev Neurosci. 2007 Oct;25(6):341–7.
81. Ioannidis JP, Munafò MR, Fusar-Poli P, Nosek BA, David SP. Publication and other reporting biases in cognitive sciences: detection, prevalence, and prevention. Trends Cogn Sci. 2014;18(5):235-41.
82. Mehta SH, Morgan JC, Sethi KD. Drug-induced movement disorders. Neurol Clin. 2015;33(1):153-74.
83. Lerner PP, Miodownik C, Lerner V. Tardive dyskinesia (syndrome): Current concept and modern approaches to its management. Psychiatry Clin Neurosci. 2015;69(6):321-34.
84. Ioannidis JP. The Mass Production of Redundant, Misleading, and Conflicted Systematic Reviews and Meta-analyses. Milbank Q. 2016;94(3):485-514.
85. Bennani YL. Drug discovery in the next decade: innovation needed ASAP. Drug Discov Today. 2011;16(17-18):779-92.
86. Sjoberg EA. Logical fallacies in animal model research. Behav Brain Funct. 2017;13(1):3.
87. Logan RW, McClung CA. Animal models of bipolar mania: The past, present and future. Neuroscience. 2016;321:163-88.
88. Neumann ID, Wegener G, Homberg JR, Cohen H, Slattery DA, Zohar J, et al. Animal models of depression and anxiety: What do they tell us about human condition? Prog Neuropsychopharmacol Biol Psychiatry. 2011;35(6):1357-75.
89. Deacon BJ. The biomedical model of mental disorder: a critical analysis of its validity, utility, and effects on psychotherapy research. Clin Psychol Rev. 2013;33(7):846-61.
90. Vollmayr B, Gass P. Learned helplessness: unique features and translational value of a cognitive depression model. Cell Tissue Res. 2013;354(1):171-8.
91. Green SB. Can animal data translate to innovations necessary for a new era of patient-centred and individualised healthcare? Bias in preclinical animal research. BMC Med Ethics. 2015;16:53.
92. Kilkenny C, Altman DG. Improving bioscience research reporting: ARRIVE-ing at a solution. Lab Anim.2010 Oct;44(4):377–8.

Figures 1 – 7

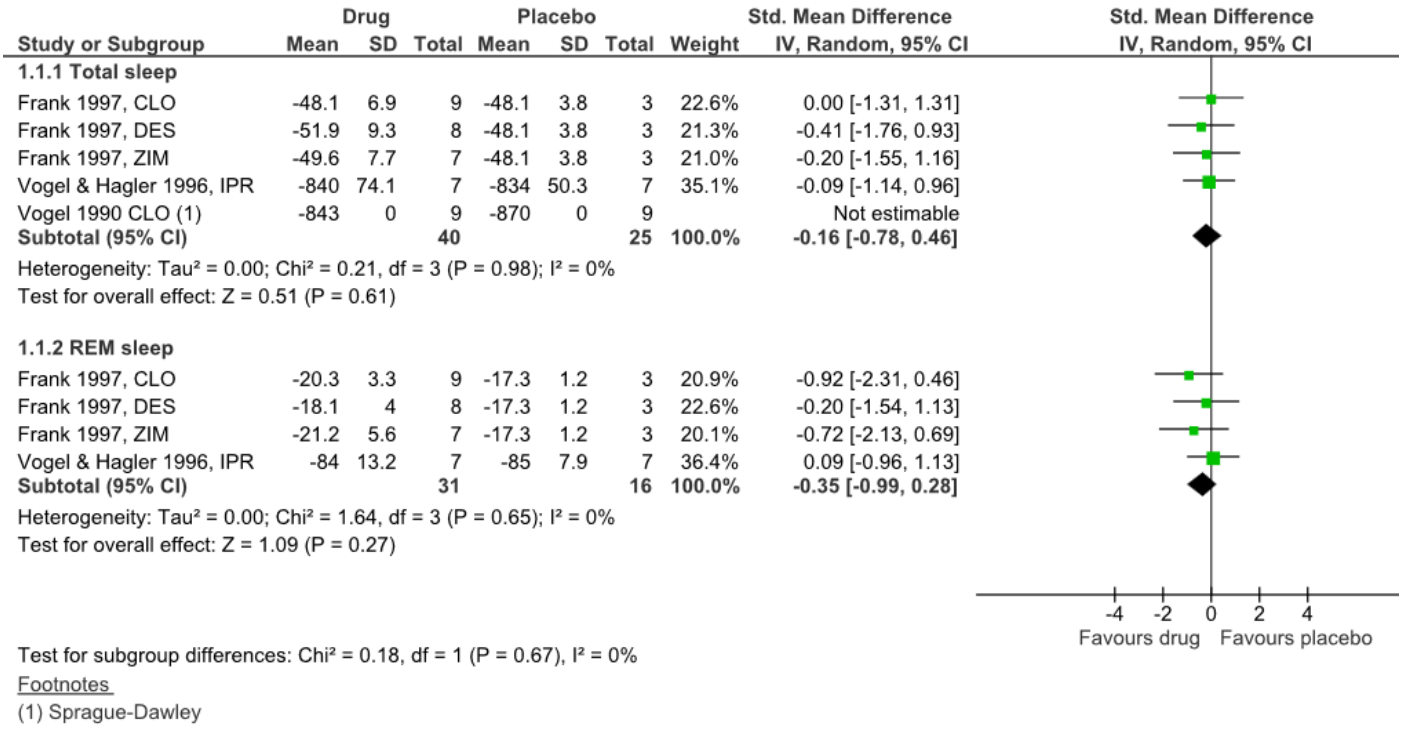


Fig 1. Studies or subgroups showing total sleep and REM sleep. Data are listed as by first author, publication year, and short for the drug used, and as hours of sleep per total time, by means with standard deviation and number of animals used (Total). IV – statistical method Inverse Variance used, CI – confidence interval, SD – standard deviation. CLO – clomipramine, DES – desipramine, ZIM – zimelidine, IPR – iprindole.

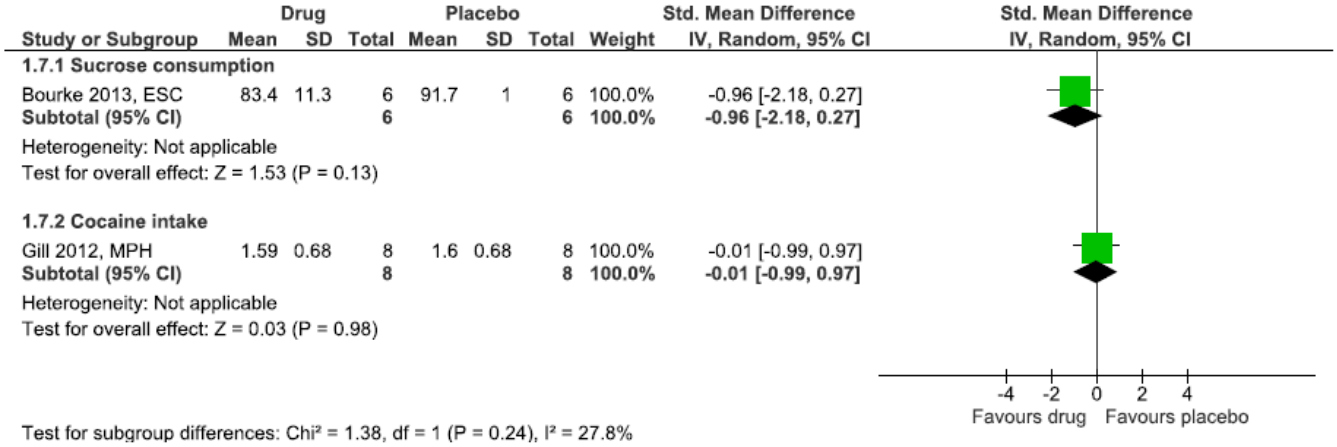


Fig 2. Studies or subgroups showing addiction. Studies are listed as by first author, publication year, and short for the drug used by means with standard deviation and number of animals used (Total). ESC – escitalopram, MPH – methylphenidate.

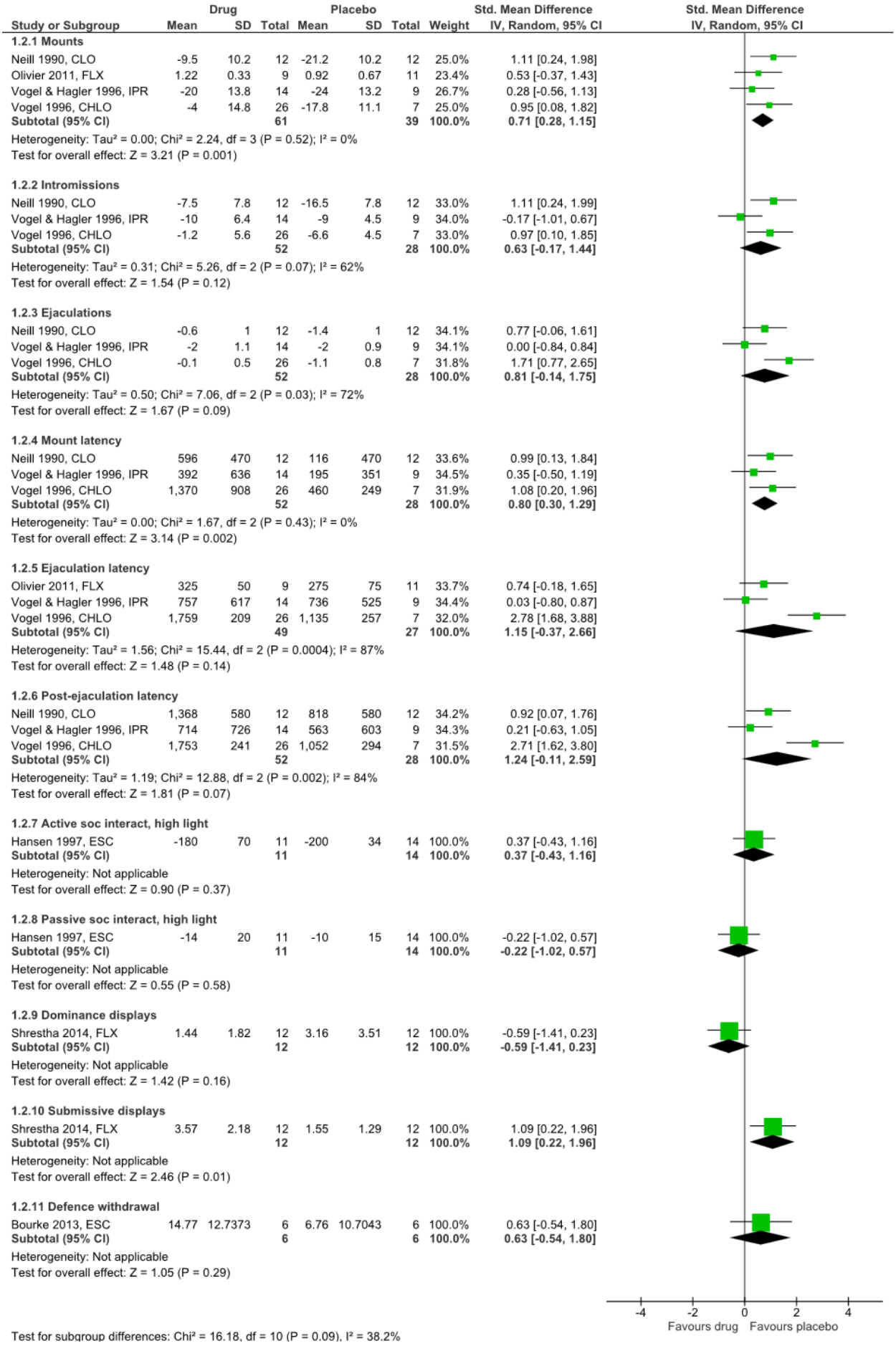


Fig 3. Studies or subgroups showing social behavior. Studies are listed as by first author, publication year, and short for the drug used, by means with standard deviation and number of animals used (Total). CLO – clomipramine, FLX – fluoxetine, IPR – iprindole, CHLO – chlorimipramine, ESC – escitalopram.

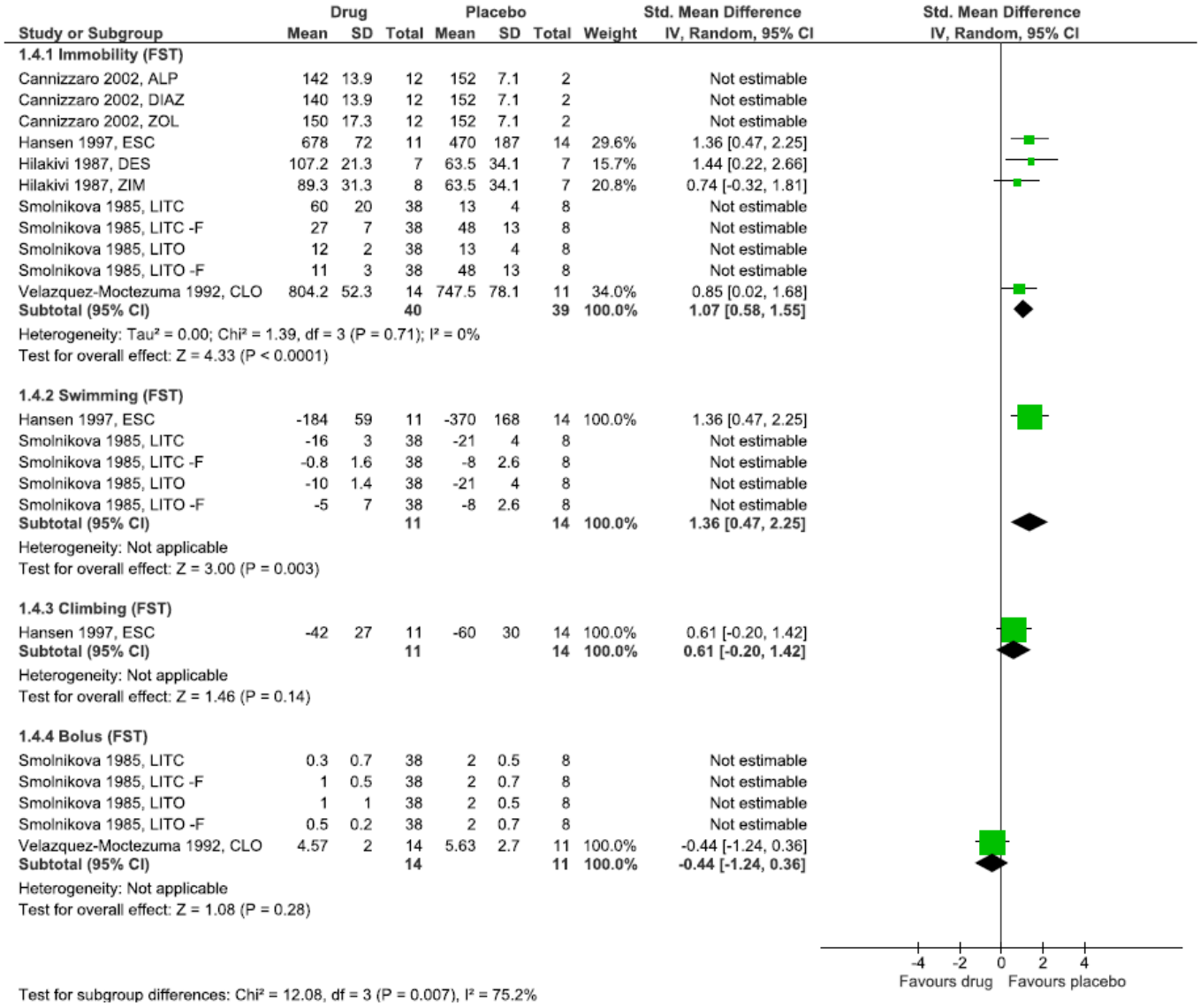


Fig 4. Studies or subgroups showing depression. Studies listed as by first author, publication year, and short for the drug used, by means with standard deviation and number of animals used (total). Suffix –f refers to the use of female animals. ALP – alprazolam, DIAZ – diazepam, ZOL – zolpidem, ESC – escitalopram, DES – desipramine, ZIM – zimelidine, LITC – lithium carbonate, LITO – lithium oxybutyrate, CLO – clomipramine.

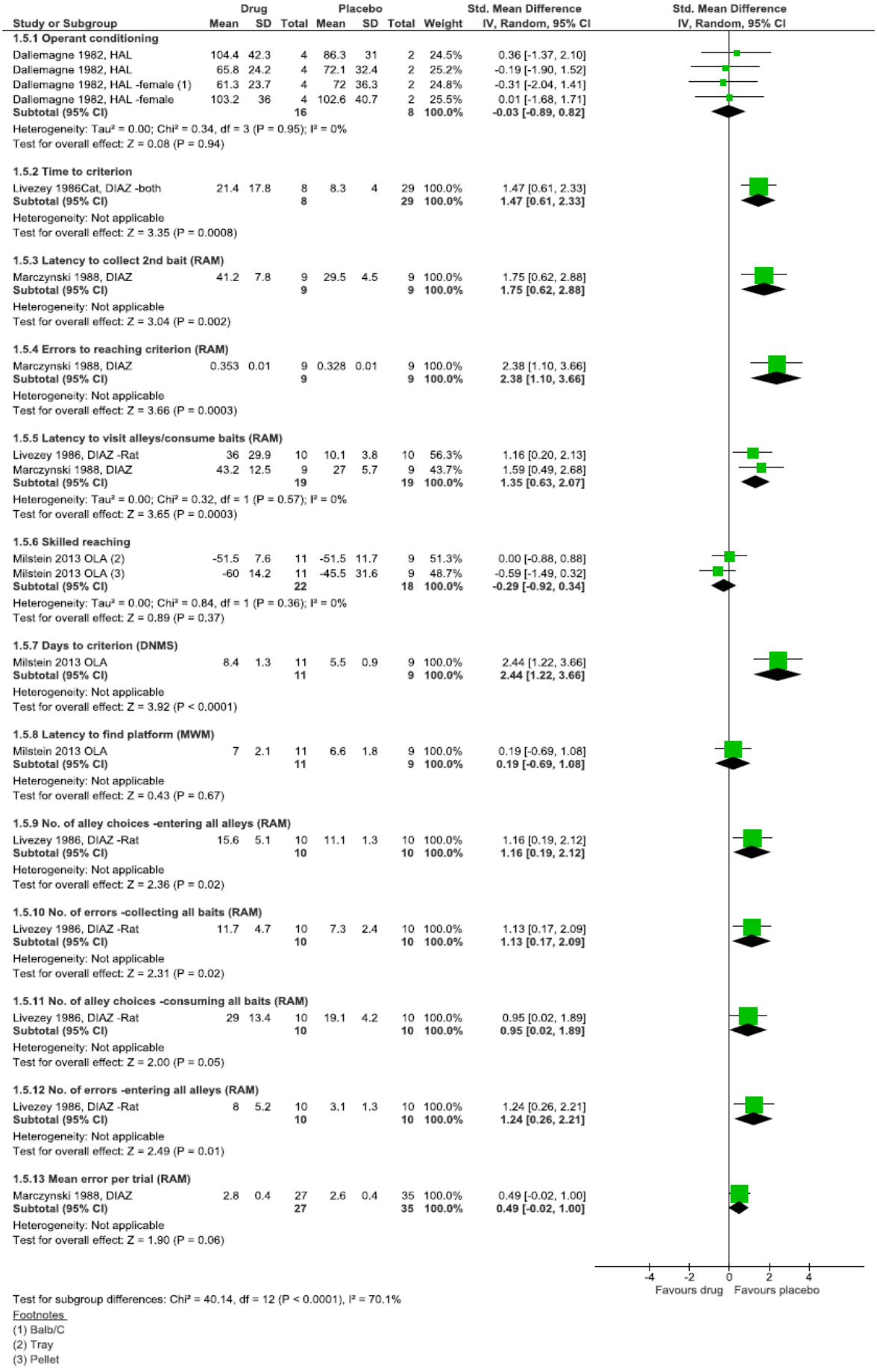
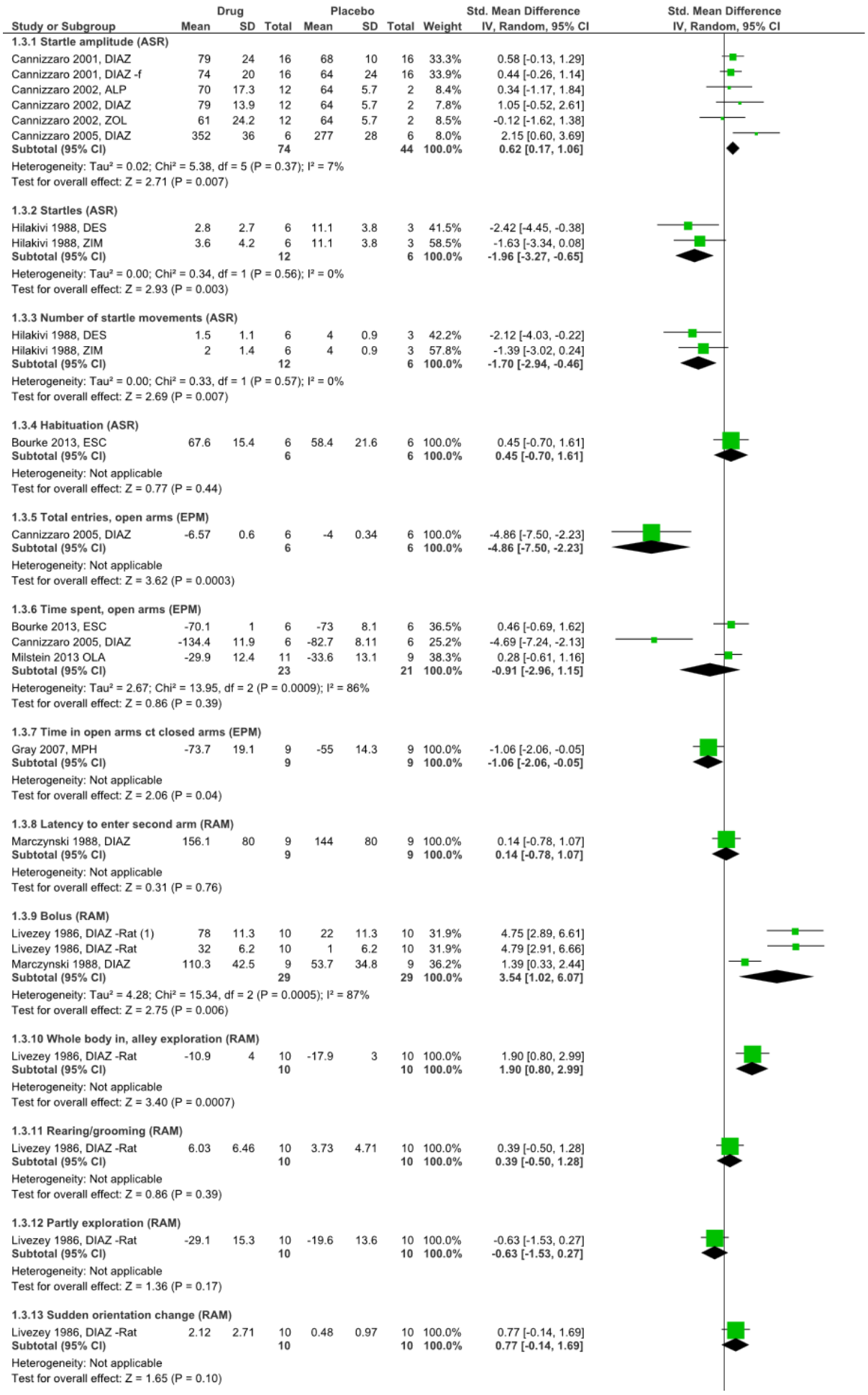


Fig 5. Studies or subgroups showing cognition, memory and learning. Studies are listed as by first author, publication year, and short for the drug used, by means with standard deviation and number of animals used (Total). Suffix –female refers to the use of female animals and suffix –both refers to the use of both female and male animals. Suffix –Rat refers to a study by Livezey et al in rats, as opposed to a study in cats. HAL – haloperidol, DIAZ – diazepam, OLA – olanzapine.


Drug
Placebo
Std. Mean Difference
Std. Mean Difference
Study or Subgroup Mean SD Total Mean SD Total Weight IV, Random, 95% CI
IV, Random, 95% CI
1.3.1 Startle amplitude (ASR)
Cannizzaro 2001, DIAZ 79 24 16 68 10 16 33.3% 0.58 [-0.13, 1.29]
Cannizzaro 2001, DIAZ -f 74 20 16 64 24 16 33.9% 0.44 [-0.26, 1.14]
Cannizzaro 2002, ALP 70 17.3 12 64 5.7 2 8.4% 0.34 [-1.17, 1.84]
Cannizzaro 2002, DIAZ 79 13.9 12 64 5.7 2 7.8% 1.05 [-0.52, 2.61]
Cannizzaro 2002, ZOL 61 24.2 12 64 5.7 2 8.5% -0.12 [-1.62, 1.38]
Cannizzaro 2005, DIAZ 352 36 6 277 28 6 8.0% 2.15 [0.60, 3.69]
Subtotal (95% CI) 74 44 100.0% 0.62 [0.17, 1.06]
Heterogeneity: Tau² = 0.02; Chi² = 5.38, df = 5 (P = 0.37); I² = 7%
Test for overall effect: Z = 2.71 (P = 0.007)
1.3.2 Startles (ASR)
Hilakivi 1988, DES 2.8 2.7 6 11.1 3.8 3 41.5% -2.42 [-4.45, -0.38]
Hilakivi 1988, ZIM 3.6 4.2 6 11.1 3.8 3 58.5% -1.63 [-3.34, 0.08]
Subtotal (95% CI) 12 6 100.0% -1.96 [-3.27, -0.65]
Heterogeneity: Tau² = 0.00; Chi² = 0.34, df = 1 (P = 0.56); I² = 0%
Test for overall effect: Z = 2.93 (P = 0.003)
1.3.3 Number of startle movements (ASR)
Hilakivi 1988, DES 1.5 1.1 6 4 0.9 3 42.2% -2.12 [-4.03, -0.22]
Hilakivi 1988, ZIM 2 1.4 6 4 0.9 3 57.8% -1.39 [-3.02, 0.24]
Subtotal (95% CI) 12 6 100.0% -1.70 [-2.94, -0.46]
Heterogeneity: Tau² = 0.00; Chi² = 0.33, df = 1 (P = 0.57); I² = 0%
Test for overall effect: Z = 2.69 (P = 0.007)
1.3.4 Habituation (ASR)
Bourke 2013, ESC 67.6 15.4 6 58.4 21.6 6 100.0% 0.45 [-0.70, 1.61]
Subtotal (95% CI) 6 6 100.0% 0.45 [-0.70, 1.61]
Heterogeneity: Not applicable
Test for overall effect: Z = 0.77 (P = 0.44)
1.3.5 Total entries, open arms (EPM)
Cannizzaro 2005, DIAZ -6.57 0.6 6 -4 0.34 6 100.0% -4.86 [-7.50, -2.23]
Subtotal (95% CI) 6 6 100.0% -4.86 [-7.50, -2.23]
Heterogeneity: Not applicable
Test for overall effect: Z = 3.62 (P = 0.0003)
1.3.6 Time spent, open arms (EPM)
Bourke 2013, ESC -70.1 1 6 -73 8.1 6 36.5% 0.46 [-0.69, 1.62]
Cannizzaro 2005, DIAZ -134.4 11.9 6 -82.7 8.11 6 25.2% -4.69 [-7.24, -2.13]
Milstein 2013 OLA -29.9 12.4 11 -33.6 13.1 9 38.3% 0.28 [-0.61, 1.16]
Subtotal (95% CI) 23 21 100.0% -0.91 [-2.96, 1.15]
Heterogeneity: Tau² = 2.67; Chi² = 13.95, df = 2 (P = 0.0009); I² = 86%
Test for overall effect: Z = 0.86 (P = 0.39)
1.3.7 Time in open arms ct closed arms (EPM)
Gray 2007, MPH -73.7 19.1 9 -55 14.3 9 100.0% -1.06 [-2.06, -0.05]
Subtotal (95% CI) 9 9 100.0% -1.06 [-2.06, -0.05]
Heterogeneity: Not applicable
Test for overall effect: Z = 2.06 (P = 0.04)
1.3.8 Latency to enter second arm (RAM)
Marczynski 1988, DIAZ 156.1 80 9 144 80 9 100.0% 0.14 [-0.78, 1.07]
Subtotal (95% CI) 9 9 100.0% 0.14 [-0.78, 1.07]
Heterogeneity: Not applicable
Test for overall effect: Z = 0.31 (P = 0.76)
1.3.9 Bolus (RAM)
Livezey 1986, DIAZ -Rat (1) 78 11.3 10 22 11.3 10 31.9% 4.75 [2.89, 6.61]
Livezey 1986, DIAZ -Rat 32 6.2 10 1 6.2 10 31.9% 4.79 [2.91, 6.66]
Marczynski 1988, DIAZ 110.3 42.5 9 53.7 34.8 9 36.2% 1.39 [0.33, 2.44]
Subtotal (95% CI) 29 29 100.0% 3.54 [1.02, 6.07]
Heterogeneity: Tau² = 4.28; Chi² = 15.34, df = 2 (P = 0.0005); I² = 87%
Test for overall effect: Z = 2.75 (P = 0.006)
1.3.10 Whole body in, alley exploration (RAM)
Livezey 1986, DIAZ -Rat -10.9 4 10 -17.9 3 10 100.0% 1.90 [0.80, 2.99]
Subtotal (95% CI) 10 10 100.0% 1.90 [0.80, 2.99]
Heterogeneity: Not applicable
Test for overall effect: Z = 3.40 (P = 0.0007)
1.3.11 Rearing/grooming (RAM)
Livezey 1986, DIAZ -Rat 6.03 6.46 10 3.73 4.71 10 100.0% 0.39 [-0.50, 1.28]
Subtotal (95% CI) 10 10 100.0% 0.39 [-0.50, 1.28]
Heterogeneity: Not applicable
Test for overall effect: Z = 0.86 (P = 0.39)
1.3.12 Partly exploration (RAM)
Livezey 1986, DIAZ -Rat -29.1 15.3 10 -19.6 13.6 10 100.0% -0.63 [-1.53, 0.27]
Subtotal (95% CI) 10 10 100.0% -0.63 [-1.53, 0.27]
Heterogeneity: Not applicable
Test for overall effect: Z = 1.36 (P = 0.17)
1.3.13 Sudden orientation change (RAM)
Livezey 1986, DIAZ -Rat 2.12 2.71 10 0.48 0.97 10 100.0% 0.77 [-0.14, 1.69]
Subtotal (95% CI) 10 10 100.0% 0.77 [-0.14, 1.69]
Heterogeneity: Not applicable
Test for overall effect: Z = 1.65 (P = 0.10)

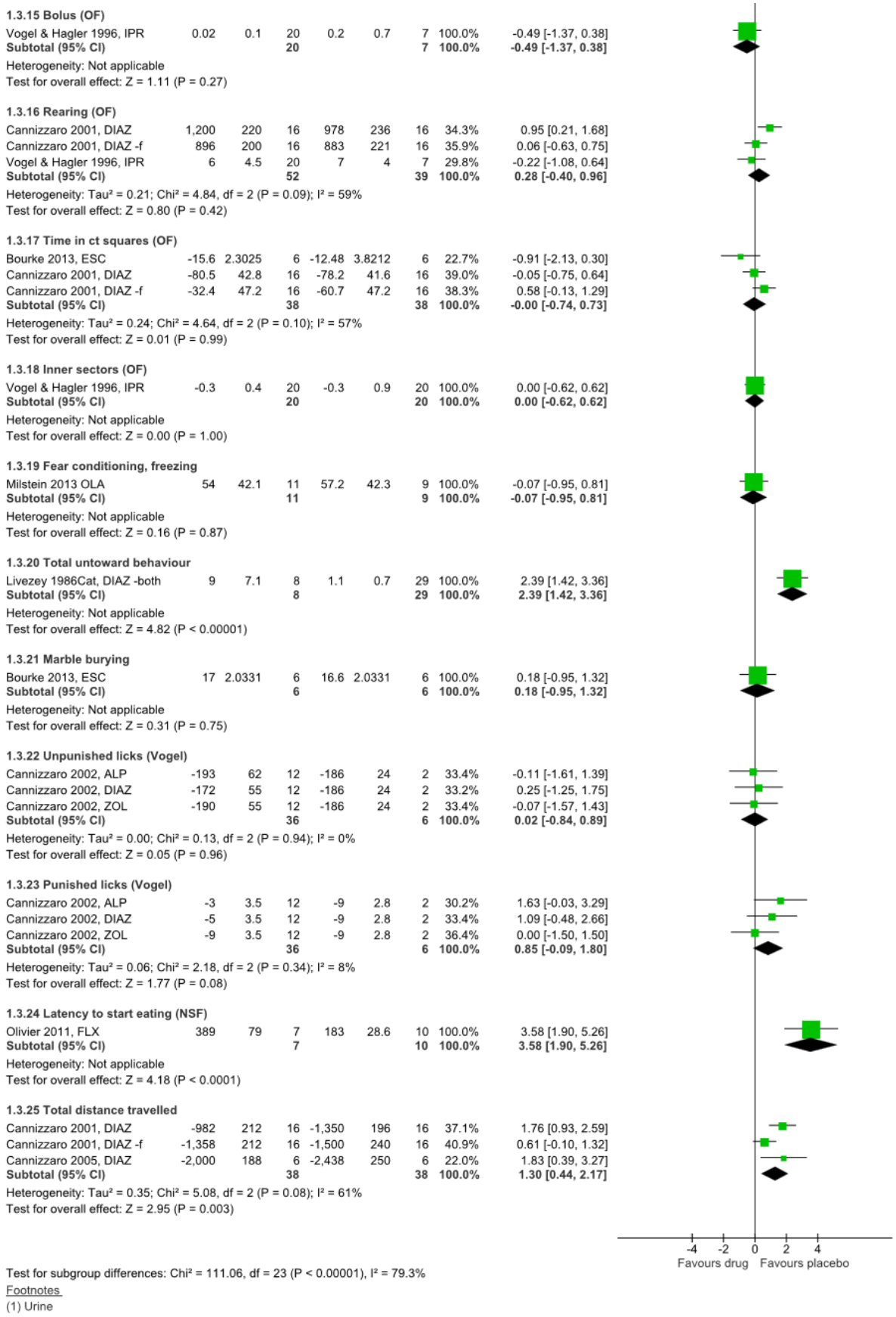


Fig 6 Studies or subgroups showing anxiety. Studies are listed by first author, publication year, and short for the drug used, by means with standard deviation and number of animals used (Total). Suffix –f refers to the use of female animals and suffix –both refers to the use of both female and male animals. Suffix –Rat refers to a study by Livezey et al in rats, as opposed to a study in cats. ALP – alprazolam, DIAZ – diazepam, ZOL – zolpidem, DES – desipramine, ZIM – zimelidine, ESC – escitalopram, OLA – olanzapine, MPH – methylphenidate, IPR – iprindole, FLX – fluoxetine.

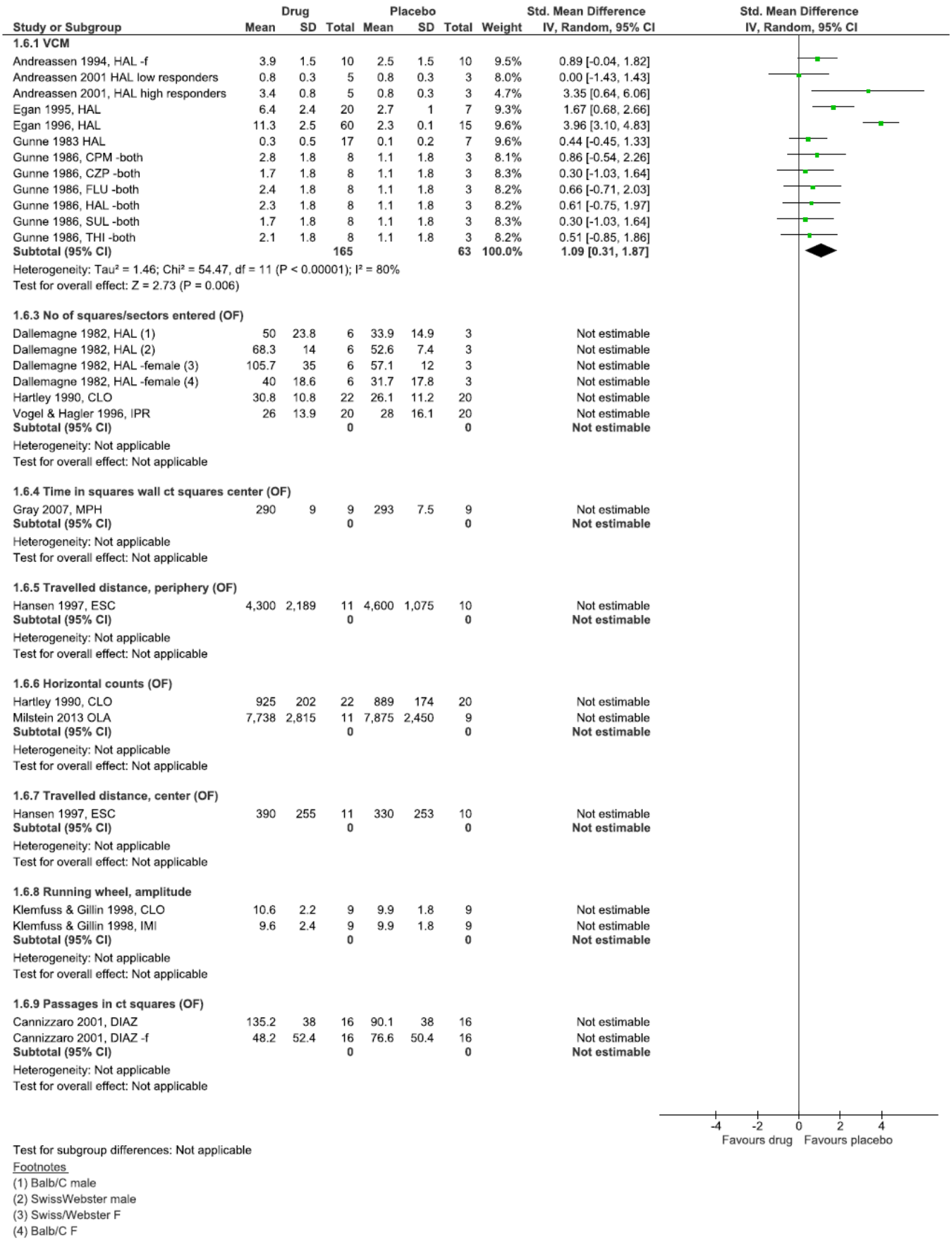


Fig 7. Subgroup analysis for antipsychotics of the outcome vacuous chewing movements (locomotion). Studies are listed by means with standard deviation and number of animals used (Total). Suffix –f refers to the use of female animals and suffix –both refers to the use of both female and male animals. HAL – haloperidol, CPM – chlorpromazine, CZP – clozapine, FLU – fluphenazine, SUL – sulpiride, THI – thioridazine.

Supporting Information S1 – short description of the tests used.

Test	Description
Vacuous chewing movements	Continuous chewing and jaw movements, despite no food present. Linked to tardive dyskinesia in humans.
Acoustic startle reflex	Used as a measure of anxiety and habituation to startle stimulus regimes.
Open field	Used to measure either locomotion or anxiety by assessing the animal's exploration level in the central area versus the outer area.
Elevated plus-maze	Used to measure either anxiety or avoidance behaviour by assessing the animal's exploration level in the open areas.
Sleep	The amount of sleep and REM sleep is measured.
Social interaction	Categories of social interaction were defined and measured. Social interaction could be sexual behaviour, passive or active, submissive or dominant or even killing behaviour, when seen in the encounter with another animal.
Forced swimming test	Used as a measure of despair or anhedonia, divided into swimming, climbing and immobile states.
Sexual behaviour	Male animals in the encounter with a female animal, outcomes are typically mounting, intromission and ejaculation and related latency periods.
Running wheel	Locomotion measured as cycles or amplitudes completed in the running wheel during a certain test time.
Radial arm maze	Used to measure working memory by placing bait in all eight arms of the maze and assess the accuracy and errors when collecting the baits.
Conditioning	Various tests were used. Classical conditioning is a process where a stimulus previously neutral comes to evoke certain behavioural responses. Operant conditioning is an extended classical conditioning where the animal behaviour is changed by rewarding or punishing a certain action.
Addiction	Measured the level of consumption of accessible stimulating compounds, e.g. sugar or recreational drugs.
Marble burying	This test uses innate rodent behaviour to assess anxiety by testing how many marbles they bury as an indication of anxiety or stress.
Vogel test	A test screening for potential anxiolytic properties of drugs. Drinking behaviour is punished with electrical shocks leading to lower water consumption in deprived animals.
Skilled reaching	Measures possible damage to motor neurons by counting the successful number of reaches for a pellet or a tray of pellets.
Morris water maze	A test for memory and learning, where the rodent has to identify the platform hidden in the non-transparent water.
Delayed non-match to sample	Measures working memory by learning choice tasks with a delay.
Tail flick test	A test for nociception, where the latency to remove the tail from a heat source is measured.
Hot plate test	A test for nociception, where the latency for the rat to either jump, use vocalization or lick the paw is measured.
Left-right alternation	A test for measuring reference and memory.

Supporting Information S2 – tests related to outcome dimensions.

Outcome	Tests included
Anxiety	Acoustic startle test, elevated plus-maze, radial arm maze, open field, conditioning, marble burying, Vogel test
Depression	Forced swimming test
Cognition, memory and learning	Conditioning, radial arm maze, skilled reaching test, delayed non-match to sample test, Morris water maze, left-right alternation
Locomotion	Vacuous chewing movements, open field, running wheel
Sleep	Sleep
Social/sexual behaviour	Sexual behaviour, social interaction
Addiction	Cocaine intake, sucrose intake
Pain	Tail flick test, hot plate test
Aggression	No studies

Supporting Information S3 – protocol.

Protocol: Long-term changes in observed behaviour after exposure to psychiatric drugs: a systematic review of animal studies

By Danborg PB, Lykkemark A, Hróbjartsson AH, Gøtzsche PC

October 2015

Background

Cases of people with psychiatric diagnoses are still increasing worldwide (1) and thus, the drugs for psychiatric disorders like bipolar disorder, mania or anxiety are still widely prescribed (2). On top of that, there is an enormous use of recently marketed drugs, approved or sold off-label, for attention deficit hyperactivity disorder (ADHD) or depression, still increasing the total amount of medication prescribed (ref). For example, sales of one group of antidepressants, SSRIs, in Denmark are now so high that every single citizen can be treated uninterruptedly for 6 years of their life (2).

Most short term studies praise the effects of the psychotropic drugs and emphasize the need for long-term assessment of efficacy (3)(4). However, data has been collected to illustrate long-term consequences, but has yet only been systematically reviewed or meta-analysed to a very small extent. Our searches only identified five systematic reviews relevant for our topic where three were performed before 1997 with human studies and case reports included (5)(6)(7).

In spring 2014, the global Cochrane Collaboration urgently called for systematic reviews on preclinical studies in order to increase translational research and transparency in clinical trials (8). It is not common practice to perform preclinical systematic reviews before conducting experiments (8) or consider bias reduction (9) as concluded by The National Centre for the Replacement, Refinement and Reduction of Animals in Research (NC3Rs, government sponsored organization in the UK) in a comprehensive study on animal studies in the UK and the US. Methodological flaws in animal studies may lead to systematic bias and thus inadequate data and wrong conclusions (10). Efficient experimental design is a foundation for high-quality research in both humans and animals and this approach will be beneficial to human science with regards to methodological approaches, minimizing research waste and of course, ultimately patient safety (11). A recent meta-analysis reported an average exaggeration of reported odds ratio (OR) of 59% on non-blinded assessments compared to blinded assessments (12). Actually, there are examples, however few, of preclinical systematic reviews revealing results that find no evidence to justify moving into clinical trials (8). From an ethical and resourcing perspective, the responsibility to conduct rigorous studies and reporting them in an accurate and systematic manner is the responsibility of the researcher for proceeding in ways that enhance the translational level and keeps the animal use at a minimal level (13).

Since the formation of Syrcle, a cooperation of researchers aiming at becoming centre of expertise on preclinical systematic reviews, new and more responsible guide lines for animal experimentation are offered to heighten the quality of research (14)(15)(16). In the light of minimizing research waste, the call for systematic reviews is even stronger. In general, there is a lack of systematic reviews and meta-analyses of adequate quality on preclinical research that will aid the selection of relevant treatment strategies in humans (17)(10), thus risk of bias must be calculated for future systematic analyses (18)(19).

A systematic review on long-term lithium therapy was inconclusive based on conflicting scientific evidence and quality of studies and reports but advised to keep lithium blood levels as low as possible (5). One study on behavioural effects of short-term lithium therapy in rodents showed lithium to be depressogenic at doses below levels that suppressed activity of the rodent (20). Lithium carbonate is used at equal levels from 1997 to 2013 (interval 20.1-23.0 x thousand packages, www.medstat.dk)

Findings from individual animal studies on antipsychotics showed that the drugs may lose their effect due to maintenance of homeostatic levels in the body and its cellular systems (21)(1); this mechanism supposedly leads to cognitive decline from long-term use of drugs.

A systematic review investigated sustained neurobehavioral effects of SSRIs in development and reported among other things lower birth weight, disturbed sleep behaviour, anhedonia and, later in life, altered sexual activity (22). There is suggestions that antipsychotic medication cause brain shrinkage and damage with long-term use (23)(24). One study investigated the effects of several antidepressants on apomorphine-induced aggression and found that all seven drugs potentiated aggressive behaviour during chronic treatment (25).

For the ADHD drug, methylphenidate, impaired growth and demineralisation of skeletal bones were seen in adolescent rats after 13 weeks of treatment; the effects was ameliorated during the five week post-treatment period (26). However, most people diagnosed with ADHD are children and adolescents growing up and thus generating bone mass. Since most children diagnosed with ADHD are boys, it is important to note that one study in rats found methylphenidate to impair normal spermatogenesis after eleven weeks of treatment, which indicated the drug to severely inflict male reproductivity (27). The use of methylphenidate for ADHD symptoms has shown to induce neuronal changes that may lead to increased drug abuse liability (28). One remarkable study showed that methylphenidate administration to young rats, corresponding in age to prepubertal children, resulted in altered levels of dopamine transporters that persisted into adulthood, long after termination of treatment (29). A study on small primates on acute administration of amphetamine derivatives showed the animals to suppress nearly all playing, eating and social interaction (30). Another study supported this depressive effect for hierarchical mid-ranking monkeys whereas for the high- and low-ranking monkeys, aggressive behaviour was increased (31).

However, the studies mentioned here are only representing a part of all articles published in the area of psychotropic drugs. There are a huge number of articles with a wide range of study designs, influencing and confounding factors and outcome measures. Furthermore, as mentioned above, the risk of bias varies greatly. Here, we wish to shed a needed light on the long-term harms of treatment with psychotropic drugs with our systematic reviews. Another Cochrane collaborator are currently investigating the effects of SSRIs on sexual behaviour in animals, thus this defined topic will be covered in an independent systematic review and not here.

Aim

As psychotropic drugs cause harms in humans, the aim of this systematic review is to investigate whether persistent changes in observed behaviour after exposure to psychotropic drugs in selected mammals also occur. We will focus on the long-term outcomes after a post-treatment period of no drugs (e.g. observed changes in social interactions, memory, cognition, mood).

Methods

Eligibility criteria

We included controlled studies on psychotropic drugs on selected healthy mammals with a post-treatment period of at least 90 days before assessment of behaviour. The reasons for a follow-up period of at least 90 days were:

- where prenatal interventions are applied, rodents are often followed up at postnatal day 90 to certainly have reached age corresponding human adulthood
- to ensure a post-treatment period of adequate length to observe changed persisting behaviours and not those that diminish after a short period of no treatment

However, the same post-treatment criteria are applied to all animals and do not distinguish between smaller and larger animals, like rodents and primates.

We excluded human studies, methodological papers and studies using animal models of disease or behaviour (e.g. animal models induced by operation, transgenic modification or drugs). Therefore, studies with a drug challenge prior to outcome assessment are also excluded, as the behavioural outcome will be affected by the drug. We also excluded studies testing non-medical substances, e.g. cannabis, 3,4-methylenedioxy-methamphetamine ('ecstasy') or nicotine as well as reporting of short-term behavioural assessments (animals followed up for less than 90 days regardless of treatment period) or non-behavioural outcomes.

Data analysis plan

We developed, through an iterative process with assistance from the professional librarians at KUB Nord, an extensive computer-based literature search on the selected subject in databases PubMed and Embase, as these are the largest and most widely used databases. The search was performed on June 2, 2014 (PubMed) using all MeSH terms in PubMed for antidepressants and antipsychotics and including also the names of the eight most sold ADHD drugs as reported by the Food and Drug Administration for 2012 (32)(atomoxetine, clonidine, dexmethylphenidate, dextroamphetamine, guanfacine, lisdexamphetamine, methylphenidate and amphetamine). A corresponding search was performed based on Map-terms in Embase on June 17, 2014, see Appendix B.

All hits (total 5647 items) provided the sample to be assessed in Zotero for possible inclusion in our systematic review and meta-analysis. The first assessment was performed by one investigator by a cautious and conservative approach on reading titles and abstracts and resulted in a total of 1091 studies (searches and from personal files). The second assessment, performed by one assessor, will confirm controlled studies on animals, where disease or behaviour is not induced by drug, transgenic modification or operation, and extract drugs tested in each study as well as drug treatment and post-treatment period of at least 90 days. The inclusion criterium regarding the post-treatment period was set to be at least 90 days, even though earlier studies and meta-analysis of preclinical studies have considered long-term to be seven days or more (33). The responsible assessor will consult a second assessor for discussion, if in doubt.

We will also write to colleagues asking them if they are aware of additional relevant studies. We will be contacting authors for outcome data not reported in the papers, where number of animals in each group is at least 15 and follow-up is at least 90 days or where data are given but clarifications of study design are needed. We will not contact authors of studies performed before 2000. Time limit to answer is approximately two months and each corresponding author will receive one reminder, if they do not answer after the first email. And finally, we will update our searches on the chosen topics for inclusion of the latest papers published, as of "today" until the day of the last search performed in PubMed and Embase.

Updated searches were performed on January 31, 2015 and revealed no further eligible studies.

Potentially eligible articles will be collected in hardcopy and read in full for third assessment (data extraction). Reference sections of obtained articles will be reviewed by reading main texts for additional papers to include due to eligibility criteria.

Data extraction (third assessment)

Two unblinded investigators will perform the data extraction independently. Extracted data will be compared and consensus reached together with a third assessor if there is any disagreement ("unblinded consensus"). More specifically, if there is any doubt about how to choose from several available data, we will ask an arbiter, who has not seen the data set.

Outcomes

AMENDED

The systematic review will analyse the observed behavioural consequences assessed in animal studies with a long post-treatment period before assessment of behaviour through various tests. We will focus on the following outcome categories, and outcomes will be grouped according to the description by the authors

- *Anxiety*
- *Depression*
- *Cognition, memory, learning*
- *Locomotion*
- *Sleep*
- *Social/sexual behaviour (encountering another animal, interaction)*
- *Addiction*
- *Aggression*

The systematic review will analyse the observed behavioural consequences assessed in animal studies with a long post-treatment period before assessment of behaviour through various tests. We will focus on the themes anxiety and depression, memory, locomotion, sleep, aggression and sexual dysfunction

The second assessment resulted in 33 studies with a follow-up of at least 90 days, regardless of sample size.

- Serial data from the follow-up period will be extracted and included (first and last time point), where there is at least one assessment after 90 days follow-up to assess whether consequences arise at day 90 or if benefits and harms diminish over time.
    - If tests are repeated on the following day, only the test from 'day 1' is reported. This is true also if there are only two test days.
- If more than one behavioural test is performed in one study, all data from behavioural tests will be extracted with a comment of whether animals are subjected to only one or more than one test.
- Lowest and highest dose for each drug will be extracted.
- We will be extracting mean end point data, not mean changes.
- If one paper reports data for multiple drugs, we will extract all data in a multiplicative manner.
- Where number of animals per group is reported as an interval, the lowest number stated will be used for analysis.
- Outcomes specified per gender will not be lumped, as most studies have used male cohorts and there can be gender differences (34)(35).
- Where multiple drug comparisons are made to one control cohort, the number of control animals will be divided by the total number of drug comparisons for the study, even though not all intervention groups are included in the analysis.
- Where individual trial data as part of a series of trials is reported, the mean and SD will be pooled.
- If various handling regimes are reported, only data for non-handled animals (those that are not handled/stroked as an intervention) will be extracted. Data on normal rearing, maternally separated animals and cross-fostered animals will be extracted.

Mean dose for all included studies are not reported in summary statistics, as doses vary by size of species, multi-dose studies and drug type.

To calculate periods of durations (e.g. post-treatment period, etc), we will calculate as follows: one full year equals 365 days, one full month equals 30 days, one week equals seven days and the rest is added in days, if periods of duration are not stated in days.

For the meta-analysis, the effect estimate is behavioural alterations and all behavioural assessments are equivalent, whereas all behavioural assessments of this outcome are included. Studies will be included for meta-analysis if the authors describe assessment of observed behaviour, per test. If there are at least two eligible studies, meta-analysis will be performed. For the meta-analysis, last time point and highest dose will be used for all drugs reported.

We will for the meta-analysis subgroup the studies to do sensitivity analysis of the impact of explanatory factors for heterogeneity. Primary factors to be assessed will be blinded observers, treatment period (as total days through the intervention period; 30 days threshold) and drug category.

All other subgroup analyses are exploratory as there are many possibly influencing factors and the interpretation of results of the subgroup analyses will be performed cautiously.

Statistical methods

For the meta-analysis, we will be reporting risk ratios for binary data, using a random effects model. For continuous data and ranking scale data, standardised mean differences (SMD; we prefer SMD (Cohen's *d*), which can combine similar but different outcome measures) will be used for analysis, inspired by Carrillo et al (33) and Wartolowska et al (36).

If there are differences in direction of scales or reporting of outcomes, then the outcomes where the higher value favours the control group will be multiplied by -1.

Pooling of means and standard deviations from several trials on the same cohort will be calculated from formula from the Cochrane Handbook for combining groups, table 7.7.a.

Outcome measures are expected to be of high heterogeneity due to the studies using various species, methods and tests during the years. Heterogeneity will be assessed as I-square (analyses outcomes relate between-group variation to total variation) and further explored by subgroup and sensitivity analyses to determine the impact of explanatory factors, as mentioned above (33)(37)(38). Meta-regression will be performed where relevant (studies n>9 for continuous data).

No data will be eliminated due to skewedness, as this might arise from the small cohorts used in several studies. Subgroup analysis will be performed for outcomes of the intervention groups where results are not skewed, i.e. ratio above 1 for probable skewedness and 2 for possible skewedness (observed mean minus lowest possible value (0) divided by SD).

Missing standard deviation will be calculated or estimated by applied methods (36), meaning that where no SD or SEM is provided, the mean SD or SEM will be calculated from stated p-values, confidence intervals or t-values, using RevMan calculator. If not applicable, SD will be imputed from the mean of three other, but similar studies (preferably those included in the analysis). To find similar studies, these points are prioritized: fitting inclusion criteria regarding non-modified species and drug class, similar test, post-treatment period and preferably more than eight animals per group.

We will not correct for multiple testing as we expect this situation to be present for the majority of included studies. For serial data, all data for the first and last time point for all outcomes will be informally compared to see if intervention group differs significantly from control group at first time point and whether the outcome scores for each group differ significantly on last time point compared to first time point.

Risk of bias will be assessed according to modified methods from Syrcle (37)(19), where the risk of bias tool from Syrcle will be used and responder selection bias is added to the bias assessment.

References

1. Whitaker R. The anatomy of an epidemic. Are psychiatric medications helpful to society? PsykoVision, Copenhagen, Denmark; 2014.
2. Deadly Medicines and Organised Crime: How Big Pharma Has Corrupted Healthcare. 1st ed. RADCLIFFE MEDICAL PRESS LTD; 2013. 310 p.
3. Godfrey J. Safety of therapeutic methylphenidate in adults: a systematic review of the evidence. J Psychopharmacol Oxf Engl. 2009 Mar;23(2):194–205.
4. Deshauer D, Moher D, Fergusson D, Moher E, Sampson M, Grimshaw J. Selective serotonin reuptake inhibitors for unipolar depression: a systematic review of classic long-term randomized controlled trials. CMAJ. 2008 May;178(10):1293–301.
5. Fountoulakis KN, Vieta E, Bouras C, Notaridis G, Giannakopoulos P, Kaprinis G, et al. A systematic review of existing data on long-term lithium therapy: neuroprotective or neurotoxic? Int J Neuropsychopharmacol Off Sci J Coll Int Neuropsychopharmacol CINP. 2008 Mar;11(2):269–87.
6. McCann UD, Seiden LS, Rubin LJ, Ricaurte GA. Brain serotonin neurotoxicity and primary pulmonary hypertension from fenfluramine and dexfenfluramine. A systematic review of the evidence. JAMA J Am Med Assoc. 1997 Aug 27;278(8):666–72.
7. Ascher JA, Cole JO, Colin JN, Feighner JP, Ferris RM, Fibiger HC, et al. Bupropion: a review of its mechanism of antidepressant activity. J Clin Psychiatry. 1995 Sep;56(9):395–401.
8. Ritskes-Hoitinga M, Leenaars M, Avey M, Rovers M, Scholten R. Systematic reviews of preclinical animal studies can make significant contributions to health care and more transparent translational medicine. Cochrane Database Syst Rev. 2014;3:ED000078.
9. Kilkenny C, Parsons N, Kadyszewski E, Festing MFW, Cuthill IC, Fry D, et al. Survey of the quality of experimental design, statistical analysis and reporting of research using animals. PloS One. 2009;4(11):e7824.
10. Van der Worp HB, Howells DW, Sena ES, Porritt MJ, Rewell S, O'Collins V, et al. Can animal models of disease reliably inform human studies? PLoS Med. 2010 Mar;7(3):e1000245.
11. Hooijmans CR, Ritskes-Hoitinga M. Progress in using systematic reviews of animal studies to improve translational research. PLoS Med. 2013;10(7):e1001482.
12. Bello S, Krogsbøll LT, Gruber J, Zhao ZJ, Fischer D, Hróbjartsson A. Lack of blinding of outcome assessors in animal model experiments implies risk of observer bias. J Clin Epidemiol. 2014 Jun 24;
13. Pound P, Bracken MB. Is animal research sufficiently evidence based to be a cornerstone of biomedical research? BMJ. 2014;348:g3387.
14. Kilkenny C, Browne W, Cuthill IC, Emerson M, Altman DG, NC3Rs Reporting Guidelines Working Group. Animal research: reporting in vivo experiments: the ARRIVE guidelines. Br J Pharmacol. 2010 Aug;160(7):1577–9.
15. Kilkenny C, Browne WJ, Cuthi I, Emerson M, Altman DG. Improving bioscience research reporting: the ARRIVE guidelines for reporting animal research. Vet Clin Pathol Am Soc Vet Clin Pathol. 2012 Mar;41(1):27–31.
16. Hooijmans CR, de Vries R, Leenaars M, Curfs J, Ritskes-Hoitinga M. Improving planning, design, reporting and scientific quality of animal experiments by using the Gold Standard Publication Checklist, in addition to the ARRIVE guidelines. Br J Pharmacol. 2011 Mar;162(6):1259–60.
17. Peters JL, Sutton AJ, Jones DR, Rushton L, Abrams KR. A systematic review of systematic reviews and meta-analyses of animal experiments with guidelines for reporting. J Environ Sci Health B. 2006;41(7):1245–58.
18. Van Luijk J, Bakker B, Rovers MM, Ritskes-Hoitinga M, de Vries RBM, Leenaars M. Systematic reviews of animal studies; missing link in translational research? PloS One. 2014;9(3):e89981.
19. Hooijmans CR, Rovers MM, de Vries RBM, Leenaars M, Ritskes-Hoitinga M, Langendam MW. SYRCLE's risk of bias tool for animal studies. BMC Med Res Methodol. 2014;14:43.
20. Tomasiewicz HC, Mague SD, Cohen BM, Carlezon, William A J. Behavioral effects of short-term administration of lithium and valproic acid in rats. Brain Res. 2006 Jun;1093(1):83–94.
21. Samaha A-N, Seeman P, Stewart J, Rajabi H, Kapur S. "Breakthrough" dopamine supersensitivity during ongoing antipsychotic treatment leads to treatment failure over time. J Neurosci. 2007 Mar;27(11):2979–86.
22. Oberlander TF, Gingrich JA, Ansorge MS. Sustained neurobehavioral effects of exposure to SSRI antidepressants during development: molecular to clinical evidence. Clin Pharmacol Ther. 2009 Dec;86(6):672–7.
23. Ho B-C, Andreasen NC, Ziebell S, Pierson R, Magnotta V. Long-term antipsychotic treatment and brain volumes: a longitudinal study of first-episode schizophrenia. Arch Gen Psychiatry. 2011 Feb;68(2):128–37.
24. Dorph-Petersen K-A, Pierri JN, Perel JM, Sun Z, Sampson AR, Lewis DA. The influence of chronic exposure to antipsychotic medications on brain size before and after tissue fixation: a comparison of haloperidol and olanzapine in macaque monkeys. Neuropsychopharmacology. 2005 Sep;30(9):1649–61.
25. Maj J, Mogilnicka E, Kordecka A. Chronic treatment with antidepressant drugs: potentiation of apomorphine-induced aggressive behaviour in rats. Neurosci Lett. 1979 Aug;13(3):337–41.

26. P T, J S, M M, H J, C S, W A, et al. Skeletal deficits and recovery in methylphenidate treated rats. Neuropsychopharmacology. 2010 Dec;
27. D K, S H, L K. The study of time dependent administration of methylphenidate on the microscopic indices of spermatogenesis and sperm analysis in adult rats. J Exp Integr Med. 2013;
28. Brandon CL, Marinelli M, White FJ. Adolescent exposure to methylphenidate alters the activity of rat midbrain dopamine neurons. Biol Psychiatry. 2003 Dec;54(12):1338–44.
29. Moll GH, Hause S, Rüther E, Rothenberger A, Huether G. Early methylphenidate administration to young rats causes a persistent reduction in the density of striatal dopamine transporters. J Child Adolesc Psychopharmacol. 2001;11(1):15–24.
30. Scraggs PR, Ridley RM. Behavioural effects of amphetamine in a small primate: relative potencies of the d- and l-isomers. Psychopharmacol Berl. 1978 Dec;59(3):243–5.
31. Smith EO, Byrd LD. Contrasting effects of d-amphetamine on affiliation and aggression in monkeys. Pharmacol Biochem Behav. 1984 Feb;20(2):255–60.
32. FDA-Approved Drugs to Treat ADHD. J Psychosoc Nurs Ment Health Serv. 2012 Mar 1;50(3):11–2.
33. Carrillo M, Ricci LA, Coppersmith GA, Melloni RH Jr. The effect of increased serotonergic neurotransmission on aggression: a critical meta-analytical review of preclinical studies. Psychopharmacology (Berl). 2009 Aug;205(3):349–68.
34. Dalla C, Pitychoutis PM, Kokras N, Papadopoulou-Daifoti Z. Sex differences in animal models of depression and antidepressant response. Basic Clin Pharmacol Toxicol. 2010 Mar;106(3):226–33.
35. Simpson J, Kelly JP. An investigation of whether there are sex differences in certain behavioural and neurochemical parameters in the rat. Behav Brain Res. 2012 Apr 1;229(1):289–300.
36. Wartolowska K, Judge A, Hopewell S, Collins GS, Dean BJF, Rombach I, et al. Use of placebo controls in the evaluation of surgery: systematic review. BMJ. 2014;348:g3253.
37. Higgins JP, Green S. Cochrane Handbook for Systematic Reviews of Interventions [Internet]. 5.1.0 ed. The Cochrane Collaboration; 2011. Available from: http://handbook.cochrane.org/
38. Vesterinen HM, Sena ES, Egan KJ, Hirst TC, Churolov L, Currie GL, et al. Meta-analysis of data from animal studies: a practical guide. J Neurosci Methods. 2014 Jan 15;221:92–102.

Appendix A: search strategies

We developed, through an iterative process with assistance from the professional librarians at KUB Nord, an extensive computer-based literature search on the selected subject in databases PubMed and Embase, as these are the largest and most widely used databases. The search was performed on June 2, 2014 (PubMed) using all MeSH terms in PubMed for antidepressants and antipsychotics and including also the names of the eight most sold ADHD drugs as reported by the Food and Drug Administration for 2012 (32)(atomoxetine, clonidine, dexmethylphenidate, dextroamphetamine, guanfacine, lisdexamphetamine, methylphenidate and amphetamine). A corresponding search was performed based on Map-terms in Embase on June 17, 2014.

The filters Comparative study, Clinical trial, Controlled clinical trial, or Randomized controlled trial was applied to reduce the number of items to screen and gave 2907 hits in PubMed. Later, a check of how many posts were included when applying each filter and revealed that filters Controlled clinical trial and Randomized controlled trial did not add any items and these two filters were thus removed from the updated search. The filter Comparative study resulted in most items (2835 items) and the filter Clinical trial resulted in 72 items. Checking the items returned when applying filter Clinical trial showed, that 16 of 20 items mentioned an animal in the title; thus, this filter will be applied in the updated searches.

The filters applied in PubMed are not applicable in Embase; this search resulted in 2704 hits with little overlap between databases, possibly due to Embase including observational studies and other studies not controlled or randomised or some studies excluded in PubMed by applying filters, but showing in Embase. Also, the data format, for which data are transferred to the reference handling program Zotero, are not similar for items identified in PubMed and Embase, thus duplicate checking are performed manually after the first screening of titles and abstracts for eligible papers.

The search in PubMed was tested against a search strategy suggested by Syrcle (based on my search presented above) with a minor difference in results (when tested by using the operator NOT and checking results, screening of 200 hits), with no differences in possibly relevant articles.

All hits (total 5647 items) provided the sample to be assessed for possible inclusion in our systematic review and meta-analysis. The first assessment was performed by one investigator by a cautious and conservative approach on reading titles and abstracts and resulted in a total of 1091 studies (searches and from personal files). The second assessment performed by one assessor will confirm controlled studies on animals where disease or behaviour is not induced by drug, transgenic modification or operation with listing of drugs tested in each study as well as drug treatment and post-treatment period of at least 90 days. The inclusion criterium regarding the post-treatment period was set to be at least 90 days, even though earlier studies and meta-analysis of preclinical studies have considered long-term to be seven days or more (33). The responsible assessor will consult a second assessor for discussion, if in doubt.

We will also write to colleagues asking them if they are aware of additional relevant studies. And finally, we will update our searches on the chosen topics for inclusion of the latest papers published, as of "today" until the day of the last search performed in PubMed and Embase.

Updated searches were performed on January 31, 2015 and revealed no further eligible papers (where filters Controlled clinical trial and Randomized controlled trial was removed).

Potentially eligible articles will be collected in hardcopy and read in full for third assessment (data extraction). Reference sections of obtained articles will be reviewed by reading main texts for additional papers to include due to eligibility criteria.

RevMan:

Studies are entered into RevMan by first author year, short for drug name (see list below).

Appendix B:

Search string, PubMed

Search	Add to builder	Query	Items found	Time
#49	Add	Search (((((((long-term) OR post-treatment) OR post-intervention) OR longitudinal) OR "Time Factors"[Mesh]) OR "Age Factors"[Mesh])) AND ((((("Monoamine Oxidase Inhibitors" [Pharmacological Action]) OR "Serotonin Uptake Inhibitors" [Pharmacological Action]) OR "Psychotropic Drugs" [Pharmacological Action]) OR (("Amphetamine"[Mesh] OR lisdexamfetamine OR "lisdexamfetamine dimesylate"[Supplementary Concept] OR "Methylphenidate"[Mesh] OR "atomoxetine"[Supplementary Concept] OR "Guanfacine"[Mesh] OR "Clonidine"[Mesh])))) AND (("non-human primate" OR "nonhuman primate" OR baboon OR macaque OR chimpanzee OR orangutan OR "rhesus monkey" OR "rhesus macaque" OR marmoset OR "cynomolgus monkey" OR gibbon OR grivet OR "squirrel monkey" OR cat OR dog OR rodent OR rat OR mouse OR rabbit OR guinea pig OR hamster OR gerbil)) Filters: Comparative Study; Clinical Trial; Controlled Clinical Trial; Randomized Controlled Trial	2907	04:13:30
#43	Add	Search (((((((long-term) OR post-treatment) OR post-intervention) OR longitudinal) OR "Time Factors"[Mesh]) OR "Age Factors"[Mesh])) AND ((((("Monoamine Oxidase Inhibitors" [Pharmacological Action]) OR "Serotonin Uptake Inhibitors" [Pharmacological Action]) OR "Psychotropic Drugs" [Pharmacological Action]) OR (("Amphetamine"[Mesh] OR lisdexamfetamine OR "lisdexamfetamine dimesylate"[Supplementary Concept] OR "Methylphenidate"[Mesh] OR "atomoxetine"[Supplementary Concept] OR "Guanfacine"[Mesh] OR "Clonidine"[Mesh])))) AND (("non-human primate" OR "nonhuman primate" OR baboon OR macaque OR chimpanzee OR orangutan OR "rhesus monkey" OR "rhesus macaque" OR marmoset OR "cynomolgus monkey" OR gibbon OR grivet OR "squirrel monkey" OR cat OR dog OR rodent OR rat OR mouse OR rabbit OR guinea pig OR hamster OR gerbil))	16128	04:13:30
#45	Add	Search (((((((long-term) OR post-treatment) OR post-intervention) OR longitudinal) OR "Time Factors"[Mesh]) OR "Age Factors"[Mesh])) AND ((((("Monoamine Oxidase Inhibitors" [Pharmacological Action]) OR "Serotonin Uptake Inhibitors" [Pharmacological Action]) OR "Psychotropic Drugs" [Pharmacological Action]) OR (("Amphetamine"[Mesh] OR lisdexamfetamine OR "lisdexamfetamine dimesylate"[Supplementary Concept] OR "Methylphenidate"[Mesh] OR "atomoxetine"[Supplementary Concept] OR "Guanfacine"[Mesh] OR "Clonidine"[Mesh])))) AND (("non-human primate" OR "nonhuman primate" OR baboon OR macaque OR chimpanzee OR orangutan OR "rhesus monkey" OR "rhesus macaque" OR marmoset OR "cynomolgus monkey" OR gibbon OR grivet OR "squirrel monkey" OR cat OR dog OR rodent OR rat OR mouse OR rabbit OR guinea pig OR hamster OR gerbil)) Filters: Meta-Analysis; Systematic Reviews	5	04:11:42
#42	Add	Search (((((long-term) OR post-treatment) OR post-intervention) OR longitudinal) OR "Time Factors"[Mesh]) OR "Age Factors"[Mesh]	1969340	04:10:31
#41	Add	Search ((("Monoamine Oxidase Inhibitors" [Pharmacological Action]) OR "Serotonin Uptake Inhibitors" [Pharmacological Action]) OR "Psychotropic Drugs" [Pharmacological Action]) OR (("Amphetamine"[Mesh] OR lisdexamfetamine OR "lisdexamfetamine dimesylate"[Supplementary Concept] OR "Methylphenidate"[Mesh] OR "atomoxetine"[Supplementary Concept] OR "Guanfacine"[Mesh] OR "Clonidine"[Mesh]))	318598	04:08:49
#40	Add	Search ("non-human primate" OR "nonhuman primate" OR baboon OR macaque OR chimpanzee OR orangutan OR "rhesus monkey" OR "rhesus macaque" OR marmoset OR "cynomolgus monkey" OR gibbon OR grivet OR "squirrel monkey" OR cat OR dog OR rodent OR rat OR mouse OR rabbit OR guinea pig OR hamster OR gerbil)	3578108	04:05:43
#39	*Add*	*Search ("Time Factors"[Mesh]) AND longitudinal*	*13203*	*04:05:09*
#38	*Add*	*Search longitudinal*	*181153*	*04:04:45*
#35	*Add*	*Search "Age Factors"[Mesh]*	*405826*	*04:04:23*
#33	*Add*	*Search "Time Factors"[Mesh]*	*985477*	*04:03:50*
#30	*Add*	*Search ("Amphetamine"[Mesh] OR lisdexamfetamine OR "lisdexamfetamine dimesylate"[Supplementary Concept] OR "Methylphenidate"[Mesh] OR "atomoxetine"[Supplementary Concept] OR "Guanfacine"[Mesh] OR "Clonidine"[Mesh])*	*35173*	*04:01:06*
#27	*Add*	*Search ("Psychotropic Drugs" [Pharmacological Action]) AND "Psychotropic Drugs"[Mesh]*	*69799*	*03:52:46*
#26	*Add*	*Search "Psychotropic Drugs" [Pharmacological Action]*	*274688*	*03:52:36*

Search	Add to builder	Query	Items found	Time
#23	Add	Search "Psychotropic Drugs"[Mesh]	127631	03:52:23
#20	Add	Search ("Serotonin Uptake Inhibitors" [Pharmacological Action]) AND "Serotonin Uptake Inhibitors"[Mesh]	15602	03:51:41
#19	Add	Search "Serotonin Uptake Inhibitors" [Pharmacological Action]	36127	03:51:33
#16	Add	Search "Serotonin Uptake Inhibitors"[Mesh]	15602	03:51:20
#12	Add	Search ("Monoamine Oxidase Inhibitors"[Mesh]) AND "Monoamine Oxidase Inhibitors" [Pharmacological Action]	8978	03:50:44
#11	Add	Search "Monoamine Oxidase Inhibitors"[Mesh]	8978	03:50:31
#8	Add	Search "Monoamine Oxidase Inhibitors" [Pharmacological Action]	19701	03:50:17
#4	Add	Search post-intervention	5410	03:25:42
#3	Add	Search post-treatment	22244	03:25:33
#1	Add	Search long-term	540769	03:24:09

Search string, EmBase

# ▲	Searches	Results
1	long-term.mp. [mp=title, abstract, subject headings, heading word, drug trade name, original title, device manufacturer, drug manufacturer, device trade name, keyword]	728823
2	post-intervention.mp. [mp=title, abstract, subject headings, heading word, drug trade name, original title, device manufacturer, drug manufacturer, device trade name, keyword]	8152
3	post-treatment.mp. [mp=title, abstract, subject headings, heading word, drug trade name, original title, device manufacturer, drug manufacturer, device trade name, keyword]	33904
4	1 or 2 or 3	767049
5	serotonin uptake inhibitor.mp.	37855
6	serotonin noradrenalin reuptake inhibitor.mp.	3069
7	monoamine oxidase inhibitor.mp.	16986
8	amphetamine/ or amphetamine plus dexamphetamine/ or amphetamine.mp.	42570
9	atomoxetine.mp.	3509
10	clonidine.mp.	38759
11	guanfacine.mp.	2538
12	methylphenidate/ or dexamphetamine/ or methylphenidate.mp.	25906
13	antidepressant agent/ or psychotropic agent/ or psychotropic.mp. or neuroleptic agent/	141911
14	5 or 6 or 7 or 8 or 9 or 10 or 11 or 12 or 13	255589
15	("non-human primate" or "nonhuman primate" or baboon or macaque or chimpanzee or orangutan or "rhesus monkey" or "rhesus macaque" or marmoset or "cynomolgus monkey" or gibbon or grivet or "squirrel monkey" or cat or dog or rodent or rat or mouse or rabbit or guinea pig or hamster or gerbil).mp. [mp=title, abstract, subject headings, heading word, drug trade name, original title, device manufacturer, drug manufacturer, device trade name, keyword]	3737433
16	**4 and 14 and 15**	**2704**

Supporting Information S4 - search strategies.

We included animals that were widely used in animal research: mice, rats, hamsters, guinea pigs, rabbits, cats, dogs, gerbils and monkeys.

Below are the applied search strategies for databases PubMed, Embase and Biosis.

PubMed search

Long-term OR post-treatment OR post-intervention OR longitudinal OR "Time Factors"[Mesh] OR "Age Factors"[Mesh]

AND

"Monoamine Oxidase Inhibitors"[Pharmacological Action] OR "Serotonin Uptake Inhibitors"[Pharmacological Action] OR "Psychotropic Drugs"[Pharmacological Action] OR "Amphetamine"[Mesh] OR lisdexamfetamine OR "lisdexamfetamine dimesylate"[Supplementary Concept] OR "Methylphenidate"[Mesh] OR "atomoxetine"[Supplementary Concept] OR "Guanfacine"[Mesh] OR "Clonidine"[Mesh]

AND

"non-human primate" OR "nonhuman primate" OR baboon OR macaque OR chimpanzee OR orangutan OR "rhesus monkey" OR "rhesus macaque" OR marmoset OR "cynomolgus monkey" OR gibbon OR grivet OR "squirrel monkey" OR cat OR dog OR rodent OR rat OR mouse OR rabbit OR guinea pig OR hamster OR gerbil

Filters applied: Comparative Study; Clinical Trial; Controlled Clinical Trial; Randomized Controlled Trial

Embase search

Long-term.mp. OR post-intervention.mp. OR post-treatment.mp.

AND

serotonin uptake inhibitor.mp. OR serotonin noradrenalin reuptake inhibitor.mp. OR monoamine oxidase inhibitor.mp. OR amphetamine/ or amphetamine plus dexamphetamine/ or amphetamine.mp. OR atomoxetine.mp. OR clonidine.mp. OR guanfacine.mp. OR methylphenidate/ or dexamphetamine/ or methylphenidate.mp. OR antidepressant agent/ or psychotropic agent/ or psychotropic.mp. or neuroleptic agent/

AND

"non-human primate" or "nonhuman primate" or baboon or macaque or chimpanzee or orangutan or "rhesus monkey" or "rhesus macaque" or marmoset or "cynomolgus monkey" or gibbon or grivet or "squirrel monkey" or cat or dog or rodent or rat or mouse or rabbit or guinea pig or hamster or gerbil).mp.

BIOSIS search

Long-term OR post-treatment OR post-intervention OR longitudinal OR "Time Factors" OR "Age Factors"

AND

"Tricyclic antidepressants" OR SSRI OR SNRI OR antidepressants OR reboxetine OR "Monoamine Oxidase Inhibitors" OR "Serotonin Uptake Inhibitors" OR "Psychotropic Drugs" OR Antipsychotics OR "Amphetamine" OR lisdexamfetamine OR "lisdexamfetamine dimesylate" OR "Methylphenidate" OR "atomoxetine" OR "Guanfacine" OR "Clonidine")

AND

"non-human primate" OR "nonhuman primate" OR baboon OR macaque OR chimpanzee OR orangutan OR "rhesus monkey" OR "rhesus macaque" OR marmoset OR "cynomolgus monkey" OR gibbon OR grivet OR "squirrel monkey" OR cat OR dog OR rodent OR rat OR mouse OR rabbit OR guinea pig OR hamster OR gerbil

Supporting Information S5 - data was extracted based on the following additional rules, besides the rules stated in the article text.

- Doses and drug variants were not pooled, each drug was reported and the highest dose was used.
- If identical tests were repeated on the following days, only the results from the first day were extracted.
- When data during follow-up were given for each consecutive day separately as part of the experimental setup (a series of connected experiments), we pooled means and SDs.
- If more than one behavioural test was performed, data from them all was extracted separately.
- If one behavioural test reported very similar outcomes for the same group of animals, the outcomes were grouped accordingly and pooled.
- Where number of animals per group was reported as an interval, the lowest number stated was used for analysis.
- Outcomes specified per gender were not to be lumped, as most studies have used male cohorts and there could be gender differences [80][81].
- If various handling regimes were reported, only data for non-handled animals was extracted.
- Data on rearing, maternally separated animals and cross-fostered animals was extracted.