N-of-1 Methods and their Contribution to Systematic Reviews and Meta-analyses

by

Salima Punja

A thesis submitted in partial fulfillment of the requirements for the degree of

Doctor of Philosophy

Experimental Medicine

Department of Medicine
University of Alberta

Abstract

Background: N-of-1 trials are prospectively planned, multiple crossover evaluations, conducted in individual patients. Evidence shows that a range of designs and statistical methods have been applied to N-of-1 trials. This thesis helps to provide a comprehensive understanding about the methodology and reporting of N-of-1 trials by synthesizing all published evidence. Furthermore, while the primary objective of N-of-1 trials is to assess treatment response in individual patients, this thesis explores whether any secondary benefits can be derived from N-of-1 trials and the data they generate. Given the number of N-of-1 trials conducted in the area of attention deficit/hyperactivity disorder, this condition was chosen as the clinical model explored in the thesis.

Objectives: i) To provide a systematic overview of published N-of-1 trials; ii) To assess how N-of-1 trials that have been conducted to assess the same interventions for the same conditions, using identical outcome measures can be aggregated in order to yield group estimates of treatment effect; and iii) To assess how N-of-1 trials can be combined with RCT data into a single meta-analysis.

Methods: A series of systematic reviews were conducted in which each review consisted of a thorough search strategy, an assessment of inclusion of primary studies, a risk of bias assessment and either a qualitative or quantitative synthesis of data. A second reviewer was involved in all reviews.

Results: This thesis found that N-of-1 trials have been conducted in over 50 conditions, and that the majority of published N-of-1 trials are published as a series. Our findings

also indicate that N-of-1 trials can be meta-analyzed across participants in order to yield population treatment effect estimates. Furthermore, we found that combining N-of-1 trials with RCT data into a single meta-analysis, impacts both the magnitude and precision of overall treatment effect estimates.

Conclusions: This thesis examined the potential for N-of-1 trials beyond their primary purpose of providing estimates of individual treatment effectiveness and demonstrates a method of aggregating N-of-1 trials across participants as well as with RCT evidence. Clinical and research recommendations on how to move this field forward have been provided.

Preface

Chapter 3 of this thesis, "Amphetamines for attention deficit/hyperactivity disorder in children and adolescents", has been accepted for publication in the *Cochrane Database of Systematic Reviews* (full citation has not yet been provided). The authors involved in this systematic review (in order as per publication) include: myself (S. Punja), L. Shamseer, L. Hartling, L. Urichuk, B. Vandermeer, C.J. Nikles and S. Vohra. I was responsible for developing the search strategies, carrying out the searches, developing the analysis plan (with the assistance of B. Vandermeer), data collection, and data analysis (under the guidance of B. Vandermeer). L. Shamseer seconded this review. All authors provided critical insight and assisted in the development of the final manuscript. S. Vohra was the supervisory author and was involved with conceiving the topic for this systematic review.

Chapter 4 of this thesis, "Amphetamines and methylphenidate for pediatric ADHD: A systematic review and meta-analysis of N-of-1 evidence", has been accepted for publication in the *Journal of Clinical Epidemiology* (full citation has not yet been provided). The authors involved in this systematic review (in order as per publication) include: myself (S. Punja), D. Xu, C.H. Schmid, L. Hartling, L. Urichuk, C.J. Nikles, and S. Vohra. I was responsible for developing the search strategies, carrying out the searches, developing the analysis plan (with the assistance of C. Schmid), data collection, and data analysis (under the guidance of C. Schmid). D. Xu seconded this review. All authors provided critical insight and assisted in the development of the final manuscript. S. Vohra was the supervisory author and was involved with conceiving the topic for this systematic review.

This body of work is dedicated to my family.

Patience ensures victory

— Hazrat Ali Ibn Abu-Talib A.S

Acknowledgements

It is a pleasure to thank the many people who made this thesis possible.

I would like to thank my supervisor, Dr. Sunita Vohra. Her constant encouragement, advice and confidence in me have been invaluable and will serve me well throughout the rest of my career. Many thanks to my supervisory committee, Dr. Lisa Hartling and Dr. Liana Urichuk, for their ongoing support and guidance.

I am especially grateful to my brother, Alim, who inspires me everyday to be better. I wish to thank my better half, Aly, for his patience and love.

Finally, I owe my deepest gratitude to my parents, who have given me the opportunity to pursue this degree and who have loved and supported me unconditionally throughout.

Table of Contents

List of Tables

List of Figures

Table of Acronyms

ADHD	Attention deficit/hyperactivity disorder
CD	Conduct disorder
CENT	CONSORT Extension for N-of-1 Trials
CONSORT	Consolidated Standards of Reporting Trials
IPD	Individual participant data
ITT	Intention to treat
MYMOP	Measure Yourself Medical Outcome Profile
ODD	Oppositional defiant disorder
PRISMA	Preferred Reporting Items for Systematic Reviews and Meta-Analyses
RCT	Randomized controlled trial
RR	Risk ratio
SD	Standard deviation
SMD	Standardized mean difference
WMD	Weighted mean difference

Chapter 1: Introduction

1.1 BACKGROUND

Evidence-based medicine is the integration of the best available evidence with clinical expertise and patient preference (1), with the ultimate goal of providing the best patient care. The parallel-group randomized controlled trial (RCT) is considered the gold-standard for determining a treatment's efficacy as it is comprised of several important features that protect against bias such as randomization, blinding, control conditions and *a priori* decisions about outcome assessment (2). There are, however, many situations in which there is a deficiency in RCT evidence or where such evidence may not be applicable to make treatment decisions. For example, challenges exist when evaluating treatments for patients with rare, unique or difficult to treat diseases, as large-scale RCTs are not available due to their prohibitive expense and difficulty achieving adequate sample sizes. Similar challenges exist with respect to pediatric research given inadequate recruitment and lack of available funding (3). Furthermore, RCTs have been criticized for their limited external validity and generalizability, since individuals with co-morbid conditions and/or patients taking concurrent therapies are often excluded (4). Restrictive eligibility criteria have been shown to limit RCT enrollment to less than 10% of individuals with the disease in question (5). Moreover, when an RCT shows a positive finding, it is often assumed that all participants improved by the same amount, when in reality not every patient derives benefit. Participants experience varying degrees of change with some doing worse, some doing better, and some staying the same. Therefore, the population treatment effect yielded by RCTs masks the important heterogeneity between participants. As a result of this lack of appropriate evidence, when treating their patients, physicians often perform what is known as a "trial of therapy." "Trials of

therapy" are utilized quite extensively in clinical practice in order to evaluate individual responses to treatment and are used for a wide variety of therapies, including medications (such as dose determination), devices, and behavioral and lifestyle therapies. In a typical trial of therapy, a patient is given the treatment in question and the subsequent clinical course determines whether or not treatment is judged to be effective and endorsed. Such informal trials are part of usual care and are unblinded, have no control group and often involve no formal outcome assessment of effectiveness. They are, therefore, particularly vulnerable to bias and can lead to false conclusions about a treatment's effectiveness. As a result, the need for an individualized approach to assessing treatment effectiveness while maintaining the rigor of an RCT is necessary in order to generate accurate and applicable treatment estimates.

The purpose of a single-subject experimental design is to allow for scientific investigation of the effectiveness of a particular treatment for an individual patient. Types of single-subject designs include: AB (baseline followed by intervention phase), ABA (baseline, intervention, return to baseline), and ABAB or N-of-1 trials (prospective multiple crossover design). The term "N-of-1 trial" is used in both the fields of medicine and behavioral sciences and has varying definitions within each. In the behavioral sciences (psychological, educational and social sciences), N-of-1 tends to be used as an umbrella term to refer to all single participant studies (including AB, ABA, and ABAB designs) (6). In medicine, however, an "N-of-1 trial" typically refers to a multiple crossover evaluation performed in a single individual whereby one arm is the treatment (A) and the other may be treatment, control or usual care (B) (7). This design can also be

described as ‘ABAB’. For the remainder of this thesis, the term “N-of-1” trial will refer to the ABAB design. N-of-1 trials use the methodological strengths of RCTs, such as randomization, blinding, the use of a control, and formalized outcome assessment. As such, N-of-1 trials minimize the risk of drawing invalid conclusions, allowing the accurate determination of treatment effectiveness for an individual. The flexibility of the N-of-1 design can be used to compare: i) an intervention versus a control (e.g. placebo, standard care); ii) two different interventions; or iii) two doses of the same intervention. Furthermore, the N-of-1 trial design can be used to compare more than 2 intervention conditions (i.e. ABCABC).

N-of-1 trials are typically conducted in individuals with chronic and relatively stable conditions (e.g. chronic pain, diabetes, attention deficit/hyperactivity disorder) (7); however, N-of-1 trials are also applicable in episodic illnesses (e.g. migraine, seizure) so long as the frequency of “attacks” is known and ideally, common [unpublished data; chapter 2 of thesis]. N-of-1 trials are not amenable for studying rapidly progressive conditions (i.e. those that are characterized by the possibility of rapid improvement or sudden catastrophic outcomes such as stroke or death). Interventions which are quick in both onset and termination of effect with modest or negligible carryover effects are most amenable to N-of-1 evaluation for practical reasons (7). Quick onset and termination diminish the need for long treatment periods and lengthy washout periods between interventions, keeping the trial short and more feasible for the patient/participant. Nevertheless, N-of-1 trials may be conducted over a longer period of time if the patient is motivated (e.g. when a disease is rare or treatment is expensive).

In an N-of-1 trial, the unit of treatment assignment (i.e. 'period') is a pre-specified time period during which the participant receives either treatment A or treatment B. The duration of the treatment period is selected to allow each treatment adequate time to take effect. A washout period may be used depending on the offset of the treatments. Treatment assignment is usually counterbalanced or randomized. The length of the trial (or number of crossovers) is usually decided *a priori*, however, a trial can continue until a clear answer is obtained (lack of benefit, clear benefit, adverse event) (7). Clinical outcomes are pre-specified and measured repeatedly over time (at least one measurement/period). The outcomes obtained during treatment A are then compared to that obtained during treatment B. These can be compared visually or statistically using a variety of tests including non-parametric tests (e.g. sign test), frequentist approaches (e.g. paired t-test) or Bayesian techniques (7, 8). These results will provide a patient with his/her individual response to treatment A versus B.

The potential advantages of N-of-1 trials are numerous and include: i) an individualized approach to assessing treatment effects for individual participants/patients without compromising methodological rigor; ii) they are ideal for assessing treatment response in participants/patients with co-morbid conditions and those using concurrent therapies; iii) they help improve patient safety by limiting therapies to only those that are demonstrated effective for a particular individual (i.e. reduction in polypharmacy); and iv) they promote personalized medicine by avoiding a 'one size fits all' approach to delivery of health care. Furthermore, given their versatility, N-of-1 trials have been successfully

applied to a range of conditions including fibromyalgia (9, 10), attention-deficit/hyperactivity disorder (11, 12), and arthritis (13, 14).

1.2 N-OF-1 TRIALS BEYOND THE INDIVIDUAL PATIENT

Although it is popularly assumed that N-of-1 trials are primarily conducted to evaluate therapeutic results in a single individual, studies have shown that the majority (55%) of published N-of-1 trials of health interventions are conducted in a series for the same condition-intervention pair (Chapter 2).

If the goal of systematic reviews is to identify and include evidence from all participants who are in an RCT through a comprehensive and exhaustive approach, then overlooking N-of-1 evidence is inconsistent with this approach. By virtue of their methods, (i.e. RCTs, albeit in a single individual), N-of-1 trials may be worthy of consideration for inclusion in systematic reviews and meta-analyses. In particular, when a series of N-of-1 trials have been conducted to assess the same intervention in similar patients, with identical outcome measures, the results may be pooled for meta-analyses in order to yield group estimates of treatment effect. This thesis aims to assess the potential secondary benefits of N-of-1 trials beyond simply assessing treatment effectiveness in individual patients, and to explore how N-of-1 trials can be aggregated across participants and across studies, as well as with RCT data in order to yield population treatment effects.

1.3 DESCRIPTION OF THE CLINICAL PROBLEM

Although the focus of this thesis is primarily methodological, the clinical condition used to answer the methodological question is pediatric attention deficit/hyperactivity disorder (ADHD). Affecting approximately 5% of children worldwide, ADHD is among the most

common pediatric psychiatric conditions (15). The condition is characterized by inattention, hyperactivity and impulsivity, which are present in two or more settings (16). The symptoms of ADHD have been shown to permeate multiple areas of functioning and have shown to have a large impact on a child's social and intellectual development (16). ADHD is associated with a number of comorbidities including anxiety, depression, oppositional defiant disorder, and conduct disorder (17, 18).

The first line of treatment for ADHD is psychostimulant medication, which includes methylphenidate and amphetamines. Evidence suggests that ADHD occurs due to insufficient production of neurotransmitters in the prefrontal cortex (19). As a result, the executive functions carried out by the prefrontal cortex are impaired (20). Psychostimulants are thought to increase levels of the neurotransmitters, norepinephrine and dopamine, in the brain by blocking their reuptake as well as increasing their production, thereby restoring executive functioning (21).

The large number of published N-of-1 trials in the area of ADHD has allowed us to explore various methodological and statistical issues around N-of-1 trials as highlighted below.

1.3 RESEARCH OBJECTIVES

The thesis is comprised of 4 chapters.

1. Chapter 2 is a systematic review of N-of-1 methods, and provides an overview of the current state of methodology and reporting in N-of-1 trials. The objective of this comprehensive review was to systematically describe: i) the study design of N-of-1 trials; ii) the statistical analyses of N-of-1 trials; and iii) the methods for combining data from a

series of N-of-1 trials.

2. Chapter 3 is a systematic review of amphetamines for pediatric ADHD based on RCT evidence. The objective of this systematic review was to assess the efficacy of amphetamines on the core symptoms of ADHD. This systematic review provided the data that contributed to chapter 5.

3. Chapter 4 includes systematic reviews of amphetamines and methylphenidate for pediatric ADHD based on N-of-1 trial evidence. The objective of these systematic reviews was to evaluate how data from N-of-1 trials may be systematically reviewed and meta-analysed by examining the effects of amphetamine and methylphenidate for attention-deficit/hyperactivity disorder (ADHD). These systematic reviews provided data that contributed to chapter 5.

4. Chapter 5 includes a combined meta-analysis of both RCT (Chapter 3) and N-of-1 (Chapter 4) data into a single meta-analysis. The objective of this meta-analysis was to assess the impact of N-of-1 trials on treatment effects in terms of magnitude and precision.

1.4 REFERENCES

1. Sackett D, Rosenberg W, Gray J, Haynes R, Richardson W. Evidence based medicine: What it is and what it isn't. *British Medical Journal* 1996; 312(7023):71-2.

2. Jadad AR. Randomized controlled trials: a user's guide. London, England: BMJ Books, 1998.

3. Gitterman DP, Hay Jr WW. That Sinking Feeling, Again? The State of National Institutes of Health Pediatric Research Funding, Fiscal Year 1992-2010. *Pediatric Research* 2008;64(5):462.

4. Fortin M, Dionne J, Pinho G, Gignac J, Almirall J, Lapointe L. Randomized controlled trials: do they have external validity for patients with multiple comorbidities? *The Annals of Family Medicine* 2006;4(2):104-108.

5. Rothwell PM. External validity of randomized controlled trials: "to whom do the results of this trial apply?" *Lancet* 2005; 365(9453):82-93.

6. Perdices M, Tate RL. Single-subject designs as a tool for evidence-based clinical practice: Are they unrecognised and undervalued? *Neuropsychological Rehabilitation* 2009;19(6):904-27.

7. Guyatt G, Jaeschke R, McGinn T. N-of-1 randomized controlled trials. In: Guyatt G, Rennie D, Meade MO, Cook DJ, editors. Users' Guides to the Medical Literature: A Manual for Evidence-Based Clinical Practice. 2nd ed. New York, NY: McGraw-Hill; 2008.

8. Zucker DR, Schmid CH, McIntosh MW, D'Agostino RB, Selker HP, Lau J. Combining single patient (N-of-1) trials to estimate population treatment effects and to evaluate individual patient responses to treatment. *Journal of Clinical Epidemiology* 1997 04;50(4):401-410.

9. Jaeschke R, Adachi J, Guyatt G, et al. Clinical usefulness of amitriptyline in fibromyalgia: the results of 23 N-of-1 randomized controlled trials. *Journal of Rheumatology* 1991;18:447–451.

10. Zucker DR, Ruthazer R, Schmid CH, et al. Lessons learned combiningN-of-1 trials to assess fibromyalgia therapies. *Journal of Rheumatology* 2006;33:2069–2077.

11. Duggan CM, Mitchell G, Nikles CJ, et al. Managing ADHD in general practice. N-of-1 trials can help! *Australian Family Physician* 2000;29:1205–1209.

12. Nikles CJ, Mitchell GK, Del Mar CB, et al. An N-of-1 trial service in clinical practice: testing the effectiveness of stimulants for attention deficit/hyperactivity

disorder. *Pediatrics* 2006;117:2040–2046.

13. March L, Irwig L, Schwarz J, et al. N-of-1 trials comparing a nonsteroidal anti-inflammatory drug with paracetamol in osteoarthritis. *British Medical Journal* 1994;309:1041–1045; discussion 1045–1046.

14. Pope JE, Prashker M, Anderson J. The efficacy and cost effectiveness of N-of-1 studies with diclofenac compared to standard treatment with nonsteroidal anti-inflammatory drugs in osteoarthritis. *Journal of Rheumatology* 2004;31:140–149.

15. Polanczyk G, Silva de Lima M, Horta BL, Biederman J, Rohde LA. Worldwide prevalence of ADHD. *American Journal of Psychiatry* 2007;4(164):942-8.

16. American Psychiatric Association. Diagnostic and Statistical Manual of Mental Disorders (DSM-IV-TR). 4th text revision edition. Washington, DC: American Psychiatric Association, 2000.

17. Agency for Health Care Policy and Research. Diagnosis of Attention-Deficit/Hyperactivity Disorder. Summary, Technical Review. http://www.ahrq.gov/clinic/epcsums/adhdsutr.htm (accessed 1 May 2013).

18. Mayes SD, Calhoun SL, Bixler EO, Vgontzas AN, Mahr F, Hillwig-Garcia J, et al. ADHD subtypes and comorbid anxiety, depression and oppositional-defiant disorder:

differences in sleep problems. *Journal of Pediatric Psychology* 2009;34(33):328-37.

19. Swanson JM, Kinsbourne M, Nigg J, et al. Etiologic subtypes of attention-deficit/hyperactivity disorder: brain imaging, molecular genetic and environmental factors and the dopamine hypothesis. *Neuropsychology Review* 2007;17(1):39-59

20. Anderson SW, Bechara A, Damasio H, Tranel D, Damasio AR. Impairment of social and moral behavior related to early damage in human prefrontal cortex. *Natural Neuroscience* 1999;2:1032-7.

21. Arnsten AFT. Stimulants: therapeutic actions in ADHD. *Neuropsychopharmacology* 2006;31:2376-83.

Chapter 2: Systematic review of the methods, statistical analysis and meta-analysis of N-of-1 trials

Salima Punja, Cecilia Bukutu, Larissa Shamseer, Margaret Sampson, Lisa Hartling, Liana Urichuk, Sunita Vohra

2.1 BACKGROUND

N-of-1 trials are prospective, multiple crossover evaluations conducted in a single-subject (i.e. ABAB) and are often randomized and blinded (1). They have a long tradition in psychological research (2) and have been used in medicine to generate treatment information when evidence from randomized controlled trials (RCTs) is not available or applicable. Three conditions should be fulfilled prior to beginning an N-of-1 trial (3). First, the condition under study should be chronic and relatively stable (e.g. autism, irritable bowel syndrome, attention deficit/hyperactivity disorder, diabetes, chronic pain). If a condition is characterized by the possibility of rapid or spontaneous improvement, such an improvement may be mistakenly attributed to the treatment under study. Second, the intervention being studied should be quick in both onset and termination of effect, therefore, mitigating the need for long treatment periods and for lengthy washout periods between interventions. Third, ideally, outcomes will be relevant to both patient and the health care provider. Disease- and patient- specific questionnaires may be used to gather data for this purpose. Standardized outcome measures can also be used when they have been validated for the condition and population under study.

Potential advantages of N-of-1 trials include: i) the approach is individualized; ii) the cost is low compared to conventional RCTs; iii) the number of people exposed to unproven therapies is minimized, but the opportunity for rigorous evidence is maintained; iv) participants will have an opportunity to experience active therapy, not just placebo; v) participants will know their results more quickly than in an RCT (e.g. months instead of years); and vi) the results will be relevant and applicable to the participants themselves. Overall, N-of-1 study design maintains methodological safeguards provided by RCTs

(blinding, randomization, controls) yet avoids the pitfalls of large trials, such as recruitment issues, prohibitive expense and lack of applicability to patients not fitting stringent eligibility criteria. Evidence-based medicine experts have suggested that the N-of-1 trial design has the potential to provide the strongest evidence for individual treatment decisions and should therefore occupy the pinnacle of the evidence pyramid (4).

Preliminary reviews reveal a range of N-of-1 designs and statistical methods in the literature (4-6). In order to optimally apply the N-of-1 methodology, all the current knowledge regarding N-of-1 trials should be synthesized. To this end, we conducted a systematic review with three objectives: i) to systematically evaluate study designs of N-of-1 trials; ii) to systematically evaluate statistical analyses of N-of-1 trials; and iii) to systematically evaluate methods for combining data (meta-analysis) from a series of N-of-1 trials. This review will provide a comprehensive understanding about the methodology and reporting of N-of-1 trials. The results of this review will lay the foundation for the development of reporting guidelines for N-of-1 trials.

2.2 METHODS

2.2.1 Search strategy

MEDLINE (1946-July week 1, 2013), Embase (1974 to 2013 Week 28), PsycInfo (1806 to July Week 2 2013), AMED (1985 to July 2013) were searched through the Ovid interface. CINAHL (from 1982, with end date unstated) was searched initially through the Ovid interface, but later through the EBSCOHost interface. Cochrane Methods Register (Issue 2 of 4, Apr 2013), Cochrane CENTRAL (Issue 6 of 12, June 2013), and the NHS Economic Evaluation Database (coverage dates unstated) were searched through

the Wiley interface. Searches were first conducted in November 2005, and updated at intervals. The most recent update search was conducted July 15-17, 2013. Reference lists of eligible studies were examined to identify additional potentially relevant studies. The full search strategy can be found in Appendices 2.10.1-2.10.7.

2.2.2 Selection criteria

English, published N-of-1 trials were selected if they met the following criteria: (i) the trial had an ABAB design [i.e. at least two interventions are compared, in which one arm is the treatment (A) and the other may be a treatment, control , usual care, or no treatment (B)]; (ii) the study contained extractable elements of design, analysis and/or meta-analysis; and (iii) the study assessed a health intervention for a particular medical/clinical condition. Studies were excluded if they followed an AB or ABA design.

2.2.3 Selection of studies

Selection of studies was based on a screening of titles and abstracts independently by two authors (SP, CB). Both reviewers independently assessed the full-text articles using the selection criteria described above. Any disagreements were resolved by a third party (SV).

2.2.4 Data extraction

One reviewer used a piloted data extraction form to extract the data and a second reviewer checked for accuracy. Extractions were done using the DistillerSR software. Extracted data included patient characteristics, treatment characteristics, design elements, methods of analysis and meta-analysis. Any disagreements were resolved through discussion.

2.2.5 Data analysis

Descriptive statistics were used to summarize all variables at the report level. Discrete variables are expressed as number and percentages whereas continuous variables are expressed as medians and ranges.

2.2.6 Unit of analysis

In this review a distinction between 'report' and 'study' is important. The former refers to the single, published entity; whereas the latter refers to the unique protocols (i.e. different participant, intervention, comparison, and outcome characteristics). One report may include multiple studies. In this review, the 'report' was used as the unit of analysis.

2.3 RESULTS

Our search yielded a total of 7394 records after duplicate removal. After screening the titles and available abstracts, 694 records were assessed for eligibility based on their full text of which 594 records were excluded. The most common reasons for exclusion included: i) not an ABAB design; ii) non-medical literature and iii) not a primary study. This left us with a total of 100 included reports (N=1995 participants), which represented 112 N-of-1 studies. The flow of reports through the screening process of the review is shown in Figure 2.9.1.

The majority (60%) of included reports were conducted as a series (i.e. one report publishing N-of-1 trial data about more than one participant for the same condition-intervention pair) and almost half came from North America. Furthermore, most reports were conducted in an outpatient setting (40%) and for research purposes (55%) as opposed to clinical evaluations (45%). The characteristics of included reports can be

found in Table 2.8.1.

The median number of enrolled participants across reports was 13 (range, 1-428) with the majority (76%) studying adult participants over 18 years of age. A wide range of conditions were evaluated using N-of-1 methodology, including diseases of the nervous system (27%), diseases of the musculoskeletal system and connective tissue (20%) and mental and behavioral disorders (17%). Table 2.8.2 presents participant characteristics of the included reports.

The median number of treatments compared across reports was 2 (range, 2-6), with the median number of treatment blocks being 3 (range, 2-15) and the median period length being 10 days (range 5 minutes to 84 days). The most common intervention studied included prescription drugs at 75%, with the most common comparison being placebo at 74%. The majority of reports only assessed two interventions (93%). Treatment characteristics of the included N-of-1 reports can be found in Table 2.8.3.

Only 21% of included reports reported a primary outcome. The most common outcome measurement tools used included Likert scales (55%), objective measures (35%; i.e. physiological assessments), visual analogue scales (30%), patient diaries (26%), and patient-generated questionnaires (18%). Only 36% of included reports addressed harms and adverse events, including their absence if that was the case. Outcome characteristics of the included reports are found in Table 2.8.4.

The design elements of the included reports can be found in Table 2.8.5. Randomization was used to determine order of treatment periods for 71% of the reports with randomization of each treatment block being the predominant unit of randomization used (55%). Blinding was also used in most of the reports (77%) with the patient/parent (74%) and clinician (43%) being the most commonly blinded parties.

Most reports (75%) statistically analyzed the individual N-of-1 trials. The most commonly used statistical methods included the paired t-test (53%) and non-parametric tests (40%). Furthermore, the most commonly reported summary measures include mean (66%) and proportion/percentage (23%), while the most commonly reported measures of precision/variance include standard deviation (24%) and 95% confidence intervals (16%). It is important to note that 49% of included papers failed to report any measure of precision/variance. See Table 2.8.6 for more characteristics of analysis of the included N-of-1 reports.

Only 37% of the 60 reports that included a series of N-of-1 trials performed a meta-analysis. The most common meta-analysis methods used include paired t-test (32%), Bayesian analysis (27%) and ANOVA (23%). The median number of participants included in the meta-analyses across these studies was 26 (range, 4-339). See Table 2.8.7 for more details on the meta-analysis of included reports.

Table 2.8.8 provides an overview of the major design and analysis challenges reported. Many researchers (70%) found it difficult to generalize their results given their small

sample sizes, sparse data, non-normal distribution, and as a result inability to perform accurate statistical analyses

2.4 DISCUSSION

The objective of this review was to provide an overview of the current N-of-1 trial literature with respect to design, analysis and meta-analysis. The results of this review indicate that there is substantial heterogeneity across these domains. Our review revealed that N-of-1 trials have been used to assess a variety of treatment options in over 50 chronic conditions, the most common being attention deficit/hyperactivity disorder, osteoarthritis, and sleep problems. Although most of the included N-of-1 trials were conducted in adults (56%), given the unique challenges that exist with respect to pediatric RCTs [i.e. underpowered and minimal available funding (7)], N-of-1 trials are garnering greater interest in pediatric research since they offer a cost-effective alternative and the benefit of rigorous evaluation in an individual patient. Furthermore, our review revealed that the majority of published N-of-1 trials are being published as a series, suggesting their value beyond simply assessing individual treatment effects and their potential to be meta-analyzed in order to provide population treatment effects which are comparable to an RCT (8). Another noteworthy finding is that the proportion of N-of-1 trials being conducted for the purposes of research and the purposes of clinical care are almost equal. This indicates the inherent value and flexibility of the N-of-1 trial in that it can be utilized as both a tool to promote evidence-based clinical care as well as a tool to produce generalizable knowledge (9).

It is clear from this review that most reports incorporated the use of elements that

maintain methodological rigour, including randomization, blinding, and formal outcome assessment. The primary goal of randomization in parallel-group RCTs is to balance both known and unknown confounders across participants in the different treatment arms. Randomization in N-of-1 trials however, is used to generate the order in which the study interventions are given to the individual. As such, the objective of randomization is slightly different than in parallel-group RCTs, as it attempts to balance known and unknown confounders over time in the single individual. Although the use of blinding in N-of-1 trials has the same objective as in group RCTs, it may play a greater role given the multiple crossover nature of the N-of-1 trial, as well as the intention of the trial to provide a clear answer on treatment response in an individual. In the event of inadequate blinding, individuals will be more likely to assess outcomes based on their pre-existing beliefs if they have the slightest idea which treatment they are receiving (10,11). This could potentially bias results, especially considering the majority of outcome measures used in N-of-1 trials are subjective. Although most studies reported formalized outcome assessment, it is important to note that only 18% utilized a patient-generated outcome questionnaire such as the Measure Yourself Medical Outcome Profile (MYMOP). The MYMOP allows the participant or patient to determine which symptoms related to their underlying condition affect them the most. These symptoms are then followed throughout the duration of the trial to track their improvement or lack thereof (12). As such, given the individualized nature of the N-of-1 trial, using a tool such as the MYMOP is particularly relevant in order to reflect what is most important to a patient.

It is evident from this review that various analysis and meta-analysis methods have been utilized for N-of-1 trials; however, research into the most appropriate methods is still

needed. There has been a push towards Bayesian analytic methods for both analyzing and meta-analyzing N-of-1 trials as this method maximizes the use of information available from each participant. The strength of the Bayesian method is that each previously conducted N-of-1 trial informs the next (13, 14). Thus, the sample size need not be defined *a priori*, and participants only need to be recruited until the study question has been answered. This allows for efficient use of resources and for questions to be answered with a minimum number of patients, making it particularly amenable for N-of-1 trials. Major barriers to conducting Bayesian analysis and meta-analysis are that it can be quite a complex model to use, and that prior information required to define parameters is often unavailable (13).

An important finding in this review is that only 44% of reports stated that they obtained ethics approval and only 53% reported that they attained informed consent. This raises an interesting point around the ethical concerns of N-of-1 trials. The ethical considerations around conducting an N-of-1 trial lay in its intent: research versus clinical care. The goal of research is to produce generalizable results, while in clinical care, the primary objective is to determine treatment effectiveness for an individual patient. This fundamental difference between the two activities, lend them to different ethical considerations (9). In clinical care, novel therapies or existing therapies for new indications may be assessed in individual patients; however, the application of these therapies is determined by clinical judgment and is therefore overseen by the usual channels for supervising clinical patient care (9). In research, however, experimental

therapies are administered by a researcher and is therefore overseen by institutional research ethics boards.

A number of barriers to conducting N-of-1 trials were reported. Researchers often found it difficult to balance study validity with feasibility. For example, when it comes to choosing the number of treatment blocks or duration of a period, it is important to have a sufficient number of treatment blocks to maintain statistical robustness, as well as long enough periods to produce observable therapeutic effects, but not too many treatment blocks or too long of periods where feasibility of the trial is jeopardized and the risk of drop-out rates increases. Furthermore, concerns over the appropriateness of the intervention/condition to be assessed by the N-of-1 trial methodology were often raised by authors. As such, it is important to determine *a priori* the suitability of the intervention of interest (i.e. does it have a quick offset and onset), for the condition of interest (i.e. is it chronic and stable), and whether or not clinically useful targets can be measured. Many reports of series of N-of-1 trials suffered from high drop-out rates which in turn resulted in an inability to conduct appropriate statistical analyses and weakened the external validity of the results. Furthermore, this lack of generalizability of results was also reported due to how participants were selected for N-of-1 trials. Since N-of-1 trials are typically conducted in individuals who are uncertain about treatment effectiveness, there may be a degree of selection bias for those who enter an N-of-1 trial and therefore may not be representative of the broader population. Conversely, some authors argued that their results had greater external validity compared to that of group RCTs given their heterogeneous participant population (i.e. inclusion of individuals with a variety of underlying comorbid conditions who are on concurrent therapies), which is far more

representative of the greater population with the condition in question. One of the most common barriers mentioned in 64% of included reports was the significant cost and time involved with carrying out N-of-1 trials. While conducting N-of-1 trials for the purposes of research may be less costly than conducting conventional RCTs, conducting N-of-1 trials for the purposes of clinical care are much more resource intensive compared to what is normally done in routine clinical care. N-of-1 trials take time to set up, implement and evaluate, while routine clinical care does not formally measure and evaluate patient outcomes; however, given that N-of-1 trials have the ability to provide more accurate assessments of treatment effectiveness in individual patients, this highlights the inadequacies of routine clinical care.

Limitations of our review are that we included only English language, published N-of-1 trials. As such, we may have overlooked a subset of N-of-1 trials that otherwise met our inclusion criteria. Furthermore, within the body of single-subject research, our definition of N-of-1 trials is relatively stringent. We examined only ABAB designs and excluded ABA and AB designs, even though these are sometimes described as N-of-1 trials. Preliminary findings show that many studies that use the ABA and AB designs are not planned investigations. While reporting is often sub-optimal, many appear to be retrospective case reports, rather than *a priori* protocols with experimental control over the two conditions (15).

It is important to note that another systematic review of N-of-1 trials has been previously published by Gabler et al (16). The key differences between the Gabler review and this

review lay in the primary objectives and therefore eligibility criteria of published reports. While this review provided a more in-depth overview around the methods utilized (e.g. characteristics of included participants; various design elements incorporated; characteristics of the statistical analysis and what was reported) in published N-of-1 trials, the Gabler review focused on how the results of N-of-1 trials impact treatment decisions in clinical care and whether published N-of-1 trials provided enough information to conduct a particular method of statistical analysis (i.e. Bayesian statistics). As such, only 60% of included reports overlapped between the two reviews. Despite these differences, the results were fairly congruent for extracted elements that were analogous to both reviews, including types of conditions studied, types of interventions assessed, length of treatment periods, number of treatment blocks utilized, and the outcome tools employed; making the results of both reviews more robust.

As is the case in group RCTs, it is clear from this review that reporting remains a major problem in N-of-1 trials. Authors failed to describe a number of elements in their reports such as ethics approval (54%), trial registration (97%), source of trial funding (69%), whether or not individuals with comorbid conditions (77%) or on concurrent therapies (69%) were included, whether or not allocation concealment was used (76%) and detection of adverse events (64%). As such, the use of a reporting guideline, such as the CONSORT statement (17), which was designed for reporting parallel group RCTs, would be useful for N-of-1 trials (18). The CONSORT Extension for N-of-1 Trials (CENT) is a comprehensive checklist designed specifically for reporting of N-of-1 trials and is currently under review. Its adoption would be helpful to improve the quality of published N-of-1 trials.

Only 3% of trials reported as having prospectively registered protocols. As such, it is likely that not all N-of-1 trials are published and readily available, particularly those that are being conducted for clinical purposes. One way of capturing these trials would be to encourage researchers and clinicians to register their N-of-1 protocols into an electronic repository (such as is done for conventional RCTs), which would also help reviewers identify selective outcome reporting and publication bias.

Although N-of-1 trials have typically been used to evaluate effectiveness of a therapy in an individual patient, this review indicates that the majority of published N-of-1 trials are conducted in a series for the same condition-intervention pair, having the potential to produce estimates of population treatment effects (8). Consequently, N-of-1 trials may be used to create predictive models to assist clinicians in more accurate prescribing to individual patients. A database of the results of conducted N-of-1 trials, both in clinical care and research, which clinicians can refer to and determine which prognostic factors match with the most successful treatment option, may result in improved patient care.

2.5 CONCLUSION

N-of-1 trials can be utilized in both clinical care and research, and have the potential to produce both individual and population treatment effect estimates. N-of-1 trials were first introduced to evidence-based medicine in the 1980's and have considerably grown since then. According to our review, there has been a 4-fold increase in the number of published N-of-1 trials over the last twenty years. This number will continue to grow as researchers, clinicians, policy-makers, and patients discover the tremendous potential of N-of-1 trials to provide patient-centered and evidence-based care.

2.6 ACKNOWLEDGEMENTS

The authors thank Drs. David Moher and Nick Barrowman for their invaluable guidance and participation in the early development of this review.

2.7 REFERENCES

1. Guyatt G, Keller J, Jaeschke R, Rosenbloom D, Adachi JD, Newhouse MT. The N-of-1
   randomized controlled trial: Clinical usefulness. Our three-year experience. *Annals of Internal Medicine* 1990; 112:293–9.

2. Kazdin AE. Single-Case Research Designs: Methods for Clinical and Applied Settings. New
   York: Oxford University Press, 1982.

3. Guyatt G, Sackett D, Adachi J et al. A clinician's guide for conducting randomized trials in
   individual patients. *Canadian Medical Association Journal* 1988; 139:497–503.

4. Guyatt G, Hayward R, Richardson WS, et al. Moving from evidence to action. In: Guyatt G,Rennie D, eds. The user's guides to the medical literature: a manual for evidence-based clinical practice. Chicago, IL: AMA Publications, 2002 p. 275–90.

5. Johannessen T, Petersen H, Kristensen P, Fosstvedt D. The controlled single subject trial. *Scandinavian Journal of Primary Health Care* 1991; 9:17–21.

6. Senn S. Suspended judgment: N-of-1 Trials. *Controlled Clinical Trials* 1993; 14(1):1–5.

7. Zarin DA, Young JL, West JC. Challenges to evidence-based medicine. *Social Psychiatry and Psychiatric Epidemiology* 2005;40(1):27-35.

8. Zucker DR, Ruthazer R, Schmid CH. Individual (N-of-1) trials can be combined to give population comparative treatment effect estimates: methodologic considerations. *Journal of Clinical Epidemiology* Dec 2010;63(12):1312-1323

9. Punja S, Eslick I, Duan N, Vohra S and the DEcIDE Methods Center N-of-1 Guidance Panel (Gabler N, Kaplan H, Kravitz R, Larson E, Pace W, Schmid C, Sim I). Design and Implementation of N-of-1 Trials: A user's guide. An Ethical Framework for N-of-1-Trials: Clinical Care, Quality Improvement, or Human Subject Research? Agency for Healthcare Research and Quality; AHRQ Publication No.13(14)-EHC122-EF. Rockville, MD. February 2014

10. Ross M, Olson JM. An expectancy-attribution model of the effects of placebos. P*sychology Review* 1981;88(5):408-37.

11. Rutherford BR, Marcus SM, Wang P, Sneed JR, Pelton G, Devanand D, Duan, Roose SP. A randomized, prospective pilot study of patient expectancy and antidepressent

outcome. *Psychological Medicine* 2013;43(5):975-82.

12. Paterson C. Measuring outcome in primary care: a patient-generated measure, MYMOP, compared to the SF-36 health survey. *British Medical Journal* 1996;312:1016-20.

13. Zucker DR, Schmid CH, McIntosh MW, D'Agostino RB, Selker HP, Lau J. Combining single patient (N-of-1) trials to estimate population treatment effects and to evaluate individual patient responses to treatment. *Journal of Clinical Epidemiology* 1997;50(4):401-410.

14. Schmid CH, Brown EN. Bayesian hierarchical models. *Methods in Enzymology* 2000;321:305-330.

15. Sampson M, Shamseer L, Bukutu C, Barrowman N, Moher D, Vohra S. Systematic reviews of N-of-1 methods, analysis and meta-analysis. Cochrane Canada 8th Annual Symposium, Ottawa: May 19-20 2010.

16. Gabler NB, Duan N, Vohra S, Kravitz RL. N-of-1 trials in the medical literature: a systematic review. *Medical Care* Aug 2011;49(8):761-768.

17. Schulz KF, Altman DG, Moher D, for the CONSORT Group. CONSORT 2010 Statement: updated guidelines for reporting parallel group randomised trials. *British*

Medical Journal 2010;340:c332.

18. Plint AC, Moher D, Morrison A, Schulz K, Altman DG, Hill C, Gaboury I. Does the CONSORT checklist improve the quality of reports of randomized controlled trials? A systematic review. *Medical Journal of Australia* 2006;185(5):263-267.

2.8 TABLES

Table 2.8.1 Study characteristics of included reports (n=100)

Characteristic	n (%)
Type of report	
Single	40 (40)
Series	60 (60)
Year of publication	
1960-1969	1 (1)
1970-1979	0 (0)
1980-1989	12 (12)
1990-1999	38 (38)
2000-2009	41 (41)
2010-2013	8 (8)
Study region	
Africa	0 (0)
Asia	1 (1)
Australia/New Zealand	17 (17)
Europe	31 (31)
North America	49 (49)
South America	2 (2)
Trial setting*	
Outpatient	40 (40)
Inpatient	17 (17)
Other	4 (4)
Not reported	42 (42)
Purpose of N-of-1 trial	
Research study	55 (55)
Clinical evaluation	45 (45)
Research ethics board status	
Approval obtained	44 (44)
Stated as not obtained	2 (2)
Not reported	54 (54)
Informed consent status	
Obtained	53 (53)
Stated as not obtained	0 (0)
Not reported	47 (47)

Trial registration	
EudraCT database	2 (2)
Clinicaltrials.gov	1 (1)
Not reported	97 (97)
Source of trial funding[†]	
Government	15 (15)
Pharmaceutical industry	15 (15)
University	4 (4)
Private/foundation	3 (3)
Not reported	69 (69)

Data are expressed as median (range) for all continuous outcomes and as n (%) for discrete variables
[*]number adds up to >100% because some reports included patients from both inpatient and outpatient settings
[†]number adds up to >100% because some reports had multiple sources of funding

Table 2.8.2 Participant characteristics of included reports (n=100)

Characteristic	**n (%)**
Total number of participants	1995
Number of enrolled participants	13 (1-428)
Number of completed N-of-1 trials	10 (1-76)[*]
Age group under study	
<18 years	16 (16)
≥18 years	76 (76)
Both	5 (5)
Not reported	3 (3)
Condition under study	
Diseases of the nervous system	Total: 27 (27)
Sleep problems/insomnia	6
Neuropathic pain	4
Cerebral palsy	3
Chronic fatigue syndrome	2
Irreversible chronic airflow limitation	2
Parkinson's disease	2
Migraine/headache	2
Chronic inflammatory demyelinating polyradiculoneuropathy	1
Epilepsy	1
Hemiparesis	1
Memory loss	1
Multiple sclerosis	1

Postherpetic neuralgia	1
Diseases of the musculoskeletal system and connective tissue	Total: 20 (20)
Osteoarthritis	8
Nonspecific pain	4
Fibromyalgia	3
Glassopharyngeal neuralgia	1
Juvenile idiopathic arthritis	1
Rheumatoid arthritis	1
Skeletal muscle cramps	1
Systemic lupus erythematosus	1
Mental and behavioral disorders	Total: 17 (17)
Attention deficit/hyperactivity disorder	6
Anxiety	2
Autism spectrum disorder	2
Schizophrenia	2
Amnesia	1
Amotivational syndrome	1
Apathy	1
Dementia	1
Depression	1
Diseases of the digestive system	Total: 11(11)
Dyspepsia	4
Nausea/vomiting	3
Chronic idiopathic intestinal obstruction	1
Gastroesophageal reflux disease	1
Oral muscositis	1
Ulcerative colitis	1
Diseases of the respiratory system	Total: 9 (9)
Asthma	2
Chronic cough	2
Chronic obstructive pulmonary disease	2
Allergic rhinitis	1
Bronchiolitis obliterans syndrome	1
Chronic airflow limitation	1
Diseases of the circulatory system	Total: 4 (4)
Hypertension	3
High international normalized ratio (at	1

risk for thrombosis)	
Endocrine, nutritional, metabolic diseases	Total: 2 (2)
Cystic Fibrosis	1
Ornithine transcarbamylase deficiency	1
Infections and parasitic diseases	Total: 2 (2)
Cryptosporidium	1
Human immunodeficiency virus	1
Other (nonspecific)	Total: 8 (8)
Abnormal electroencephalograms	1
Akinetic mutism	1
Dyspnea	1
Gustatory facial sweating	1
Hand injury	1
Inflammation of continent ileostomy	1
Nocturia	1
Range of conditions (not specified)	1
Diagnostic criteria used to confirm condition[†]	
Formal diagnostic criteria	31 (31)
Clinical judgment	16 (16)
Other	1 (1)
Not reported	55 (55)
Inclusion of patients/participants with comorbidities	
Yes	23 (23)
No	0 (0)
Not reported	77 (77)
Inclusion of patients/participants on concurrent therapies	
Yes	29 (29)
No	2 (2)
Not reported/unclear	69 (69)

Data are expressed as median (range) for all continuous outcomes and as n (%) for discrete variables

[*]only 79/100 reports reported on the number of completers

[†]number adds up to >100% because some reports used multiple diagnostic criteria

Table 2.8.3 Treatment characteristics of included reports (n=100)

Characteristic	
No. of treatments compared	2 (2-6)
Number of repeated cycles/treatment blocks	3 (2-15)
Period length (in days)*	10 (5 minutes to 84)
Interventions under study	
Prescription drug	75 (75)
Natural health product	13 (13)
Device	5 (5)
Behavioral, cognitive	2 (2)
Other	5 (5)
Comparison/control studied†	
Placebo	74 (74)
Active	25 (25)
A different dose	3 (3)
Usual care	2 (2)
No treatment	8 (8)
Washout period	
Yes-intervention-free period	18 (18)
Yes-analytic washout	20 (20)
No	38 (38)
Not reported	24 (24)
Reasons for early stopping of treatments‡	
Lack of efficacy	5 (21)
Adverse event	13 (55)
Marked improvement	2 (8)
Participant condition worsened	2 (8)
Other	2 (8)

Data are expressed as median (range) for all continuous outcomes and as n (%) for discrete variables

*6 reports did not report their period length

†adds up to >100 since some reports assessed more than one comparison

‡early stopping was reported 24 times

Table 2.8.4 Characteristics of outcomes of included reports (n=100)

Characteristic	n (%)
Primary outcome reported	
Yes	21 (21)
Not reported	79 (79)
Selection of the primary outcome(s) (n=21)*	
Patient/Parent	5 (24)
Investigator	10 (48)
Clinician	1 (5)
Not reported	9 (43)
Outcome measurement tools used†	
Likert Scale	55 (55)
Objective measure (e.g. blood pressure)	35 (35)
Visual analogue scale	30 (30)
Diary	26 (26)
Patient generated questionnaire (e.g. MYMOP)	18 (18)
Other	7 (7)
Not reported	1 (1)
Assessment of harms	
Yes	36 (36)
Not reported	64 (64)

Data are expressed as median (range) for all continuous outcomes and as n (%) for discrete variables

*numbers add up to >100% since some primary outcomes were chosen collaboratively by multiple parties

† number adds up to >100% since some reports used multiple outcome measurement tools

Table 2.8.5 Design characteristics of included reports (n=100)

Characteristic	
Sample size calculation*	
Yes	7 (12)
No	4 (7)
Not reported	49 (82)
Run-in period	
Yes	27 (27)
No	7 (7)
Not reported	66 (66)
Length of run-in (n=27; in days)	22 (1-90)
Reason for run-in (n=27)†	
Establish tolerable dose	12 (44)
Compliance with study protocol	3 (11)
Washout	2 (7)
Baseline data	1 (4)
Other	3 (11)
Not reported	7 (26)
Randomization	
Yes	71 (71)
No	4 (4)
Not reported	25 (25)
Method used to generate randomization sequence (n=71)	
Computer-generated	12 (17)
Random number table	5 (7)
Coin toss	3 (4)
Other	1 (1)
Not reported	50 (70)
Unit of randomization used (n=71)	
Within each treatment block	39 (55)
Entire sequence	6 (8)
By period	3 (4)
Not reported	23 (32)
Allocation concealment	
Yes	22 (22)

No	2 (2)
Not reported	76 (76)
Blinding	
Yes	77 (77)
No	2 (2)
Stated as not possible for the intervention	2 (2)
Not reported	19 (19)
Blinded parties (n=77)[‡]	
Patient/Parent	57 (74)
Clinician	33 (43)
Investigator	21 (27)
Outcome assessor	13 (17)
Research assistant	3 (4)
None reported	18 (23)

Data are expressed as median (range) for all continuous outcomes and as n (%) for discrete variables

[*]Refers only to reports of series of N-of-1 trials (n=60)

[†]Adds up to >100% because some reports had >1 reason for run-in

[‡]Adds up to >100% because some reports blinded multiple parties

Table 2.8.6 Characteristics of analysis of included reports (n=100)

Characteristic	n (%)
Analysis of individual N-of-1 trials[*]	
Statistically	75 (75)
Visually/graphically	14 (14)
Not reported	18 (18)
Method of statistical analysis (n=75)[†]	
Paired t-test	40 (53)
Bayesian	5 (7)
ANOVA (regression)	3 (4)
Non-parametric test	
Wilcoxon rank sum rest	11 (15)
Mann-Whitney	3 (4)
Sign test	3 (4)
Kruskal-Wallis	2 (3)
McNemar's test	1 (1)
Other	4 (5)
Other	8 (11)
Not reported/unclear	7 (9)
Summary measures reported[‡]	

Mean	66 (66)
Proportion/percentage	23 (23)
Median	8 (8)
Endpoint or change scores/period	4 (4)
Frequency of events/period	3 (3)
Other	2 (2)
None reported	11 (11)
Measures of precision reported[§]	
Standard deviation	24 (24)
95% confidence interval	16 (16)
Standard error	9 (9)
Interquartile range	5 (5)
90% confidence interval	3 (3)
None reported	49 (49)
Assessment of carryover effect	
Yes	9 (9)
Not reported	91 (91)
Detection of carryover effect (n=9)	
Yes	3 (33)
No	6 (67)
Assessment of period effect	
Yes	3 (3)
Not reported/unclear	97 (97)
Detection of period effect (n=3)	
Yes	2 (67)
No	1 (33)

Data are expressed as median (range) for all continuous outcomes and as n (%) for discrete variables

[*]number adds up to >100% since some reports used both statistical and graphical methods to assess treatment effect

[†]number adds up to >100% since some reports used more than one method of analysis

[‡]number adds up to >100% since some reports reported more than 1 summary measures

[§]number adds up to >100% since some reports reported more than 1 measure of precision

Table 2.8.7 Characteristics of meta-analysis of included reports that include a series of N-of-1 trials (n=60)

Characteristic	
Meta-analysis conducted	
Yes	22 (37)
No	34 (57)
Not reported	4 (6)
Number of participants included in the meta-analysis (n=22)	26 (4-339)
Method used to pool the data (n=22)	
Paired t-test	7 (32)
Bayesian	6 (27)
ANOVA	5 (23)
ANCOVA	1 (4)
Not reported	3 (14)
Additional subgroup/sensitivity analyses conducted (n=22)	
Yes	3 (14)
Not reported	19 (86)

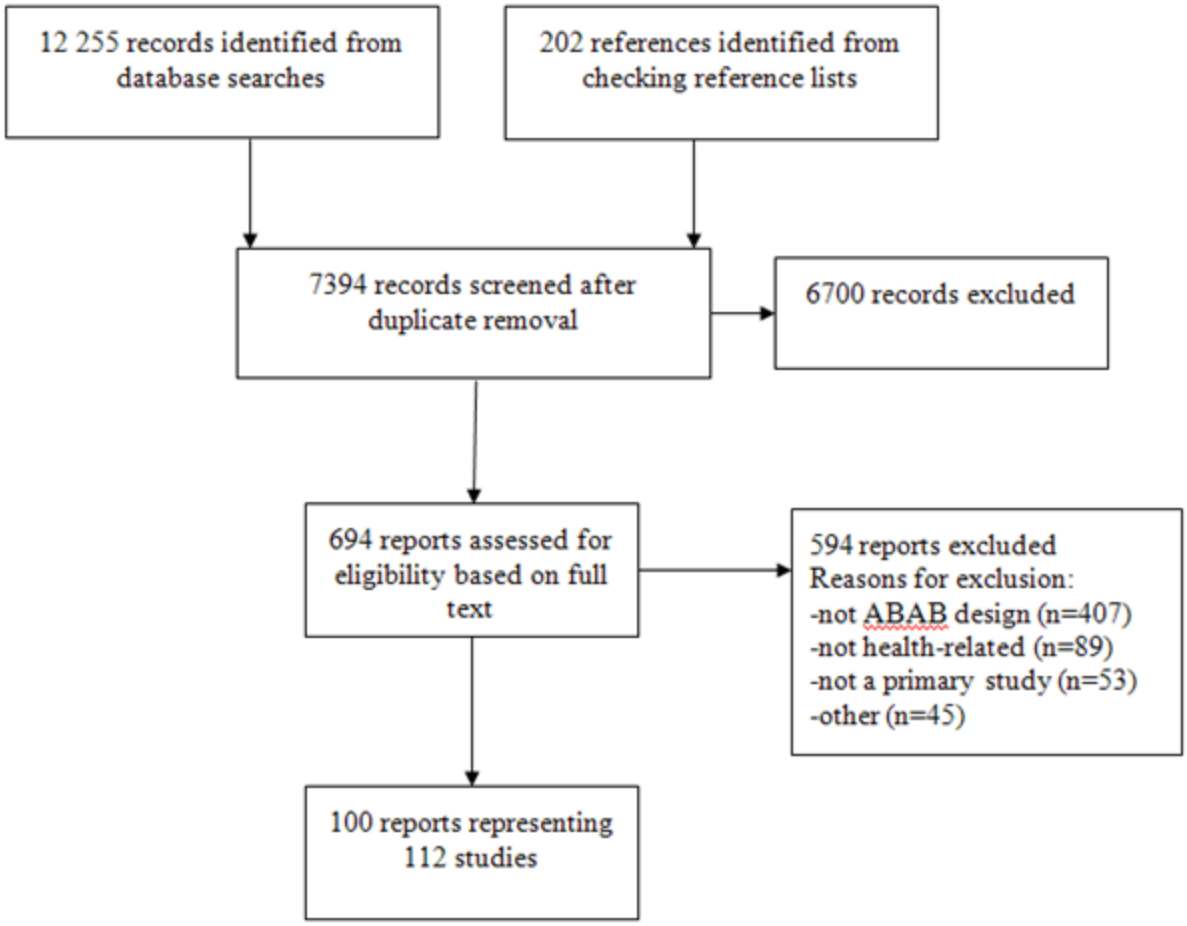
Data are expressed as median (range) for all continuous outcomes and as n (%) for discrete variables

Table 2.8.8 Most commonly reported limitations/barriers of included reports

Design
Too few treatment blocks
Short treatment periods
High drop-out rates
Small sample sizes
Too few measurements
Inappropriate methodology for intervention/condition under study
Inability to blind due to type of intervention
Analysis
Large amounts of missing data made statistical testing inappropriate
Non-normal distribution of data
Limited power of statistical tests with too few treatment blocks
Other
Cost and time of conducting an N-of-1 trial

2.9 Figures

Figure 2.9.1 Flow of studies



2.10 Appendices

2.10.1 MEDLINE search strategy

1. N-of-1.tw.
2. (individual$ adj2 trial$).tw.
3. IMET$.tw.
4. or/1-3
5. Double-blind method/
6. Research Design/
7. Randomized Controlled Trials/
8. Random Allocation/
9. Clinical Trials/
10. Models, Statistical/
11. Cross-Over Studies/

12. Placebos/
13. (Bayes$ or frequentist).mp.
14. or/5-13
15. mahon jl.au.
16. guyatt g$.au.
17. Feldman BM.au.
18. johannessen t$.au.
19. or/15-18
20. 4 and (14 or 19)
21. n of 1.ti.
22. abab.ti,ab.
23. (single adj (subject or patient or case) adj3 (trial$ or design)).tw.
24. (n of 1 adj3 (trial$ or rct$ or random$ or challenge$)).tw.
25. ((series or random$ or multiple) adj3 n of 1).tw.
26. n of 1 service$.tw.
27. individuali#ed medication effectiveness test$.tw.
28. patient$ as their own control$.tw.
29. or/21-27
30. ("10028447" or "10088595" or "10366668" or "10534600" or "10573255" or "10573256" or "10591309" or "10616901" or "10692633" or "10796613" or "10832027" or "10836333" or "11127076" or "11140235" or "11180571" or "11202717" or "11316255" or "11347795" or "11456506" or "11487728" or "11562563" or "1158877" or "11603973" or "11769218" or "12147558" or "12147570" or "12422252" or "12482471" or "12649627" or "12650412" or "1290621" or "14221" or "14644852" or "14659636" or "14705210" or "14705233" or "1494653" or "14969791" or "15009785" or "15022126" or "15022653" or "1502480" or "15052922" or "15142921" or "1517652" or "1518731" or "15203041" or "1520842" or "15546268" or "15578800" or "15590502" or "15603678" or "1562961" or "15662295" or "15662296" or "15673992" or "15722401" or "15743108" or "1575078" or "15808030" or "15808031" or "15808032" or "15911473" or "15914517" or "1592840" or "1592952" or "15950705" or "15965209" or "15971593" or "15984899" or "16052316" or "16118811" or "16279260" or "1663776" or "1663849" or "1663858" or "1675018" or "1697199" or "1724672" or "1757405" or "1808614" or "1856813" or "1928975" or "2139111" or "2161315" or "2194767" or "2265191" or "2297206" or "2304823" or "2309473" or "2309475" or "2405971" or "2457686" or "2502209" or "2811648" or "2876383" or "2956558" or "3144178" or "3201140" or "3263146" or "3327883" or "3409132" or "3409138" or "3410840" or "3621904" or "3704110" or "37248" or "3795040" or "3991824" or "6326609" or "6927687" or "7503169" or "7560704" or "7703779" or "7703780" or "7715318" or "7716695" or "7819919" or "7843782" or "7850886" or "7854793" or "7861872" or "7874096" or "7893278" or "7918449" or "7945494" or "7950736" or "7994389" or "8033025" or "8058148" or "8089690" or "8263770" or "8289103" or "8380194" or "8440091" or "8507664" or "8518976" or "8616414" or "8624326" or "8747973" or "8761252" or "8761253" or "8807426" or "8822684" or "8927783" or "8939323" or "8961779" or "8987134" or "9000744" or "9179098" or "9257324" or "9310903" or "9311056" or "9341168" or "9527698" or "9601160" or "9679365" or "9684421" or "9712612" or "9732386" or "9789474" or "9842829" or "9925061").ui.

31. or/20,28,30
32. 30 not (animals/ not humans/)
33. 32 not (acta crystallographica or Spectrochimica or Spectrometry or Physical Review or Molecul* or chemical or chemistry or organic letters).jw.

Note: Line 30 represents reviewer nominations from checking reference lists.

2.10.2 Embase search strategy

1. "1".ti. /freq=2
2. n.ti. /freq=2
3. N-of-1.tw.
4. (individual$ adj2 trial$).tw.
5. IMET$.tw.
6. or/3-5
7. Double Blind Procedure/
8. Methodology/
9. Randomized Controlled Trials/
10. Randomization/
11. Clinical Trial/
12. Statistical Model/
13. Crossover Procedure/
14. Placebo/
15. (Bayes$ or frequentist).mp.
16. or/7-15
17. mahon jl.au.
18. guyatt g$.au.
19. Feldman BM.au.
20. johannessen t$.au.
21. or/17-20
22. 6 and (16 or 21)
23. n of 1.ti.
24. abab.ti,ab.
25. (single adj (subject or patient or case) adj3 (trial$ or design)).tw.
26. (n of 1 adj3 (trial$ or rct$ or random$ or challenge$)).tw.
27. ((series or random$ or multiple) adj3 n of 1).tw.
28. n of 1 service$.tw.
29. individuali#ed medication effectiveness test$.tw.
30. patient$ as their own control$.tw.
31. or/21-28
32. or/22,31
33. limit 32 to human
34. 33 not (acta crystallographica or Spectrochimica or Spectrometry or Physical Review or Molecul* or chemical or chemistry or organic letters).jw.
35. 34 not (1 or 2)
36. ((N-of-1 adj "1*") or (N-of-1 adj "2*") or ((N-of-1 adj "3*") or (N-of-1 adj "4*") or (N-of-1 adj "5*") or (N-of-1 adj "6*") or (N-of-1 adj "7*") or (N-of-1 adj "8*")) or (N-of-1 adj "9*") or (N-of-1 adj "0*")).ti,ab.

37. "n = 1*".ab.
38. "N-acetyltransferase".ti.
39. 36 or 37 or 38
41. 35 not 40

2.10.3 PsychInfo search strategy

1. N-of-1.tw.
2. (individual$ adj2 trial$).tw.
3. IMET$.tw.
4. or/1-3
5. Double blind.mp.
6. Quasi Experimental Methods/
7. Experimental Design/
8. Random$.mp.
9. Treatment Effectiveness Evaluation/
10. Statistical Analysis/ or Mathematical Modeling/
11. (Cross-over or crossover).mp.
12. Placebo/
13. (Bayes$ or frequentist).mp.
14. or/5-13
15. mahon jl.au.
16. guyatt g$.au.
17. Feldman BM.au.
18. johannessen t$.au.
19. or/15-18
20. 4 and (14 or 19)
21. n of 1.ti.
22. abab.ti,ab.
23. (single adj (subject or patient or case) adj3 (trial$ or design)).tw.
24. (n of 1 adj3 (trial$ or rct$ or random$ or challenge$)).tw.
25. ((series or random$ or multiple) adj3 n of 1).tw.
26. n of 1 service$.tw.
27. individuali#ed medication effectiveness test$.tw.
28. patient$ as their own control$.tw.
29. or/21-28
30. limit 29 to human
31. limit 29 to animal
32. 29 not (31 not 30)
33. ((N-of-1 adj "1*") or (N-of-1 adj "2*") or ((N-of-1 adj "3*") or (N-of-1 adj "4*") or (N-of-1 adj "5*") or (N-of-1 adj "6*") or (N-of-1 adj "7*") or (N-of-1 adj "8*")) or (N-of-1 adj "9*") or (N-of-1 adj "0*")).ti,ab.
34. 32 not 33

2.10.4 AMED search strategy

1. N-of-1.tw.
2. (individual$ adj2 trial$).tw.

3. IMET$.tw.
4. or/1-3
5. Double-blind method/
6. Research Design/
7. Randomized Controlled Trials/
8. Random Allocation/
9. Clinical Trials/
10. (Cross-over or crossover).mp.
11. Placebos/
12. (Bayes$ or frequentist).mp.
13. or/5-12
14. mahon jl.au.
15. guyatt g$.au.
16. Feldman BM.au.
17. johannessen t$.au.
18. or/14-17
19. 4 and (13 or 18)
20. n of 1.ti.
21. abab.ti,ab.
22. (single adj (subject or patient or case) adj3 (trial$ or design)).tw.
23. (n of 1 adj3 (trial$ or rct$ or random$ or challenge$)).tw.
24. ((series or random$ or multiple) adj3 n of 1).tw.
25. n of 1 service$.tw.
26. individuali#ed medication effectiveness test$.tw.
27. patient$ as their own control$.tw.
28. or/20-27
29. or/19,28
30. ((N-of-1 adj "1*") or (N-of-1 adj "2*") or ((N-of-1 adj "3*") or (N-of-1 adj "4*") or (N-of-1 adj "5*") or (N-of-1 adj "6*") or (N-of-1 adj "7*") or (N-of-1 adj "8*")) or (N-of-1 adj "9*") or (N-of-1 adj "0*")).ti,ab.
31. 29 not 30
33. 31 and 32

2.10.5 CINAHL search strategy

S27 S15 or S26
S26 S16 OR S17 OR S18 OR S19 OR S20 OR S21 OR S22 OR S23 OR S24 OR S25
S25 patient* as their own control*
S24 individuali?ed medication effectiveness test*
S23 series N3 n of 1 OR Random* N3 n of 1 OR multiple N3 n of 1
S22 n of 1 W3 trial* OR n of 1 W3 rct* OR n of 1 W3 random* OR n of 1 W3 challenge*
S21 (single W1 participant) N3 design* OR (single W1 participant) N3 design*
S20 (single W1 subject) N3 design* OR (single W1 subject) N3 design*
S19 (single W1 case) N3 trial* OR (single W1 case) N3 trial*
S18 TX (single W1 patient) N3 trial* OR (single W1 patient) N3 design*
S17 ABAB
S16 TI n of 1

S15 S4 and S14
S14 S5 OR S6 OR S7 OR S8 OR S9 OR S10 OR S11 OR S12 OR S13
S13 Bayes* or frequentist
S12 MH Placebos
S11 MH Crossover Design
S10 MH "Models, Statistical"
S9 MH Clinical Trials
S8 Random*
S7 MH Quasi-Experimental Studies
S6 MH Study Design
S5 MH Double-Blind Studies
S4 (S1 OR S2 OR S3)
S3 TX IMET*
S2 TX individual* N2 trial*
S1 TX N-of-1

2.10.6 Cochrane Database (CENTRAL, Methods Registry, NHS Economic Evaluation Database) search strategy

1. N-of-1":ti,ab,kw
2. "n of 1".ti,ab,kw
3. #1 or #2

2.10.7 ERIC and Sociological Abstracts (2007 update only) search strategy

N-of-1 or (n of 1) or (single subject) or (individual* within 3 trial*) or IMET* or ABAB

Chapter 3: Amphetamines for attention deficit/hyperactivity disorder in children and adolescents

Salima Punja, Larissa Shamseer, Lisa Hartling, Liana Urichuk, Ben Vandermeer, Catherine J Nikles, Sunita Vohra

Accepted by the *Cochrane Database of Systematic Reviews*

3.1 ABSTRACT

Background: Attention deficit/hyperactivity disorder (ADHD) is among the most common psychiatric conditions affecting children and adolescents. Amphetamines are among the most commonly prescribed medications to manage ADHD. There are three main classes of amphetamines: dexamphetamine, lisdexamphetamine and mixed amphetamine salts, which can be further broken down into short- and long-acting formulations. A systematic review assessing their efficacy and safety in this population has never been conducted

Objectives: To assess the efficacy and safety of amphetamines for ADHD in children and adolescents. Search methods: In July 2013 we searched CENTRAL, PubMed, Embase, PsycInfo, ProQuest Dissertation and Theses and Networked Digital Library of Theses and Dissertations. We also searched ClinicalTrials.gov, and checked the reference lists of relevant studies and reviews identified by the searches.

Selection criteria: Parallel and crossover randomized controlled trials comparing amphetamine derivatives against placebo in a pediatric population (<18 years) with ADHD.

Data collection and analysis: Two authors independently extracted data on participants, settings, interventions, methodology and outcomes for each included study. For continuous outcomes, we calculated standardized mean difference (SMD) and for dichotomous outcomes we calculated relative risk (RR). The most commonly reported adverse events in the primary studies were meta-analysed. Data was meta-analysed using a random-effects model.

Results: We included 20 trials with 2183 participants. Study durations ranged from 14 to

365 days, with the majority lasting less than 6 months. Seven studies were parallel-group trials, while 13 studies were crossover. Amphetamines significantly improved total ADHD core symptom severity according to parent ratings (SMD -0.40, 95% CI -0.53 to -0.27), teacher ratings (SMD -0.55, 95% CI -0.83 to -0.27) and clinician ratings (SMD -0.84, 95% CI -1.32 to -0.36). In addition, the proportion of responders was significantly higher when participants were taking amphetamines as compared to placebo (RR 3.12, 95% CI 2.32 to 4.20).

The most commonly reported adverse events included decreased appetite, insomnia, abdominal pain, headaches, anxiety, and nausea/vomiting. Amphetamines were associated with a significantly higher proportion of participants experiencing decreased appetite (RR 7.44, 95% CI 2.99 to 18.48), insomnia (RR 3.67, 95% CI 1.83 to 7.38) and abdominal pain (RR 1.65, 95% CI 1.17 to 2.31). In addition the proportion of participants who experienced at least one adverse event was significantly higher in the amphetamine group (RR 1.31, 95% CI 1.16 to 1.48).

Subgroup analyses were performed for amphetamine preparation (dexamphetamine, lisdexamphetamine, mixed amphetamine salts), amphetamine release formulation (long-acting versus short-acting), and funding source (industry versus non-industry). No statistically significant between-group differences were observed in any of the subgroups across any of the outcomes.

Conclusion: Although amphetamines were effective at reducing the core symptoms of ADHD in the short-term, they were associated with a number of adverse events. This review showed no evidence that supports any one amphetamine derivative over another, and does not reveal any differences between long-acting and short-acting amphetamine

preparations. Future research should be longer in duration (i.e. >12 months), include more qualitative outcomes (e.g. quality of life and parent stress), and be transparently reported.

3.2 PLAIN LANGUAGE SUMMARY

Attention deficit/hyperactivity disorder (ADHD) is a common problem affecting children and adolescents. ADHD is characterized by inattention, impulsivity and hyperactivity. One of the most popular treatments for managing ADHD is amphetamines, which are a class of stimulant medications. In order to assess the effects of amphetamines, we searched for clinical trials of children with ADHD. We found 20 studies which together indicate that amphetamines are effective at improving the symptoms of ADHD. The trials also indicate, however, that amphetamines are linked to a number of adverse effects. As such, when deciding on a course of treatment, the risks and benefits must be weighed against each other. Further research is needed to determine the long-term effects (i.e. >12 months) of amphetamines in children with ADHD.

3.3 BACKGROUND

3.3.1 Description of the condition

Attention deficit hyperactivity disorder (ADHD) is among the most common pediatric psychiatric conditions, affecting around 5% of children worldwide (1). ADHD is characterized by three core symptoms: inattention, impulsivity, and hyperactivity, which are more frequently displayed than would be typical in children of the same age (2). The core symptoms are often presented to various degrees in different children, breaking

ADHD down into three subtypes: i) the predominantly inattentive type; ii) the predominantly hyperactive-impulsive type; and iii) the combined type (i.e. children displaying both inattention and hyperactivity) (2). The condition is often diagnosed through a rigorous set of criteria at a young age, usually between the ages of three and six years old (3). The potential for comorbidities is extremely high in this population and they are present in almost two-thirds of pediatric ADHD cases, with the most common being oppositional defiant disorder (ODD) (50%), conduct disorder (CD) (35%), anxiety disorder (33%), and depression (33%) (4, 5).

The symptoms of ADHD have been shown to permeate a child's performance across multiple settings, having long-term effects on their academic performance and social development. Studies have also shown that children with ADHD are more likely to be irritable, impatient, and aggressive (6). In addition, families who have children with ADHD often experience higher levels of parental stress and frustration, marital disruption, and social isolation (7). It has been estimated that approximately 65% of childhood ADHD cases will persist into adulthood (8), making it a chronic lifetime condition for many.

3.3.2 Description of the intervention

A wide variety of treatments have been used for the management of ADHD including psychosocial interventions, dietary management, herbal and homeopathic remedies, and biofeedback; however, for the past few decades, the psychostimulant, methylphenidate, has been the first line of treatment (2) and has been found to be effective in 70% to 90% of school-aged children (6, 9). Amphetamines are the second most frequently prescribed

psychostimulants for pediatric ADHD, and are becoming an increasingly popular alternative for children who fail to respond to methylphenidate (10). There are currently three different amphetamine preparations available, including: 1) dexamphetamine (dextroamphetamine or d-amphetamine sulfate), which comes in both short-acting formulations (e.g. Dextrostat, Dexedrine) and long-acting formulations (e.g. Dexedrine Spansules, Dexedrine SR); 2) lisdexamphetamine, which is available as a long-acting formulation (e.g. Vyvanase); and 3) mixed amphetamine salts, which also comes in both short-acting (e.g. Adderall) as well as long-acting preparations (e.g. Adderall XR) (10, 11).

3.3.3 How the intervention might work

Although the pathophysiology of ADHD is poorly understood, evidence has suggested that ADHD may be the result of insufficient production of norepinephrine and dopamine in the prefrontal cortex (12). As such, the executive functions carried out by the prefrontal cortex are impaired, resulting in forgetfulness, distractibility, impulsivity, and inappropriate social behaviours (13). Others believe that the limbic system plays a major role in the pathophysiology of ADHD, and it is thought that hyperactivity and impulsivity result from abnormally low tonic dopamine activity within this region of the brain (14). In either case, as a psychostimulant, amphetamines are thought to both promote marked neurotransmitter release into the synaptic cleft as well as disrupt normal reuptake of neurotransmitters thereby increasing levels of norepinephrine and dopamine in these regions of the brain and resulting in improved executive functioning (15). A Cochrane review of amphetamines for ADHD in adults found they improved short-term symptom severity (16).

3.3.4 Why it is important to conduct this review

Despite being one of the most thoroughly researched disorders in medicine, one of the major controversies regarding ADHD is the use of psychostimulants as a treatment option. While current evidence suggests that amphetamines may be beneficial for improving the core symptoms of ADHD, their effects on academic and social domains remain inconsistent and unclear (6). A wide variation in the use and prescription of amphetamines across communities suggests that there is a lack of consensus among practitioners regarding which patients with ADHD should be treated with amphetamines. Charach et al. (17) and Miller et al. (18) have conducted reviews assessing amphetamines for pediatric ADHD; however, the former only focused on long-term effectiveness of amphetamines (i.e. >12 months), while the latter is not only out of date, it focused solely on the dexamphetamine preparation. As primary stakeholders, it is imperative for healthcare providers, parents, and those diagnosed with ADHD to be aware of the most suitable treatment options available, and how they differ in terms of their efficacy and safety profiles. Our synthesis of all available randomized controlled trials of the efficacy and safety of amphetamines for pediatric ADHD will provide evidence to better inform clinical practice and further research relating to ADHD management. While assessing amphetamines against other ADHD treatments such as methylphenidate, psychotherapy and antidepressants is important, establishing whether amphetamines are superior to placebo is a necessary first step, thus this review will focus only on the amphetamine versus placebo comparison.

3.4 OBJECTIVES

To assess the efficacy and safety of amphetamines for ADHD in children and adolescents.

3.5 METHODS

3.5.1 Criteria for considering studies for this review

3.5.1.1 Types of studies

Parallel and crossover randomized controlled trials (RCTs).

3.5.1.2 Types of participants

Children and adolescents (less than 18 years of age) with a clinical diagnosis of ADHD according to specified diagnostic criteria, such as the DSM-III (19), DSM-IV (20), or equivalent (note: since the DSM 5 was released during the conduct of this review, studies utilizing this criteria are not included in this review). We included trials that involved participants with some comorbid conditions (oppositional defiant disorder (ODD), conduct disorder (CD), anxiety, depression). We excluded trials that included participants with psychiatric comorbidities, which require highly specialized treatment programs (for example, autism, bipolar disorder, psychosis).

3.5.1.3 Types of interventions

Intervention: any oral form of amphetamine (i.e. dexamphetamine, lisdexamphetamine, mixed amphetamine salts), at any dose.

Control: placebo.

3.5.1.4 Types of outcome measures

Primary outcomes

1. Change in core ADHD symptoms (inattention, hyperactivity, impulsivity), as measured by a validated scale rated by children, parents, teachers, clinicians, or investigator

Secondary outcomes

1. Clinical improvement measured the by the Clinical Global Impression-Improvement (CGI-I) scale*
2. Academic performance*
3. Parental stress
4. Quality of life*
5. Retention: proportion of randomized participants who completed the trial

Adverse events

1. Proportion of adverse events
2. Proportion of participants who experienced at least one adverse event as reported in the trials*
3. Proportion of participants who withdrew due to any adverse event

Outcomes marked with an * were used to populate the 'Summary of findings' table.
No data were available on the outcome 'Parental Stress'.

3.5.2 Search methods for identification of studies

3.5.2.1 Electronic searches

We searched the following electronic databases on July 9, 2013.

1. Cochrane Central Register of Controlled Trials (CENTRAL) (Ovid platform)
2. Medline (Ovid platform; 1948- July 9, 2013)
3. EMBASE (Ovid platform; 1974- July 9, 2013)

4. PsycINFO (Ovid platform; 1806- July 9, 2013)
5. ClinicalTrials.gov (http://clinicaltrials.gov/)
6. ProQuest Dissertations and Theses
7. Networked Digital Library of Theses and Dissertations

No language or date restrictions were applied.

The full search strategy can be found in Appendices 3.13.1-3.13.4.

3.5.2.2 Searching other resources

We inspected the reference lists of identified RCTs and reviewed articles to identify additional publications.

3.5.3 Data collection and analysis

3.5.3.1 Selection of studies

Two review authors (SP and LS) independently screened all the titles and abstracts retrieved from the search to identify those that appeared to meet the inclusion criteria. We obtained the full-text articles of those studies and assessed their eligibility. Disagreements were resolved by SV.

3.5.3.2 Data extraction and management

Two review authors (SP and LS) independently extracted data related to study methods, participant characteristics, and outcomes by using a predesigned data collection form. Disagreements were resolved through discussion. All the relevant data was entered into Review Manager (21) by SP.

3.5.3.3 Assessment of risk of bias in included studies

For each included study, two review authors (SP and LS) independently assessed the risk of bias for the seven domains explained below using the Cochrane 'Risk of bias' tool (22). Disagreements were resolved through discussion.

The following sources of bias were assessed as being at low risk of bias, high risk of bias, or unclear risk of bias:

Random sequence generation

Description: the method used to generate the allocation sequence is described in sufficient detail so as to assess whether it should have produced comparable groups.

Review authors' judgement: what is the risk of selection bias due to inadequate generation of a randomized sequence?

Allocation concealment

Description: the method used to conceal the allocation sequence is described in sufficient detail to determine whether intervention allocations could have been foreseen in advance of, or during, enrolment.

Review authors' judgement: what is the risk of selection bias due to inadequate concealment of allocations prior to assignment?

Blinding of participants and personnel

Description: measures used, if any, to blind study participants and personnel from knowledge of which intervention a participant received and any information relating to whether the intended blinding was effective.

Review authors' judgement: what is the risk of performance bias due to knowledge of the allocated interventions by participants and personnel during the study?

Blinding of outcome assessment

Description: measures used, if any, to blind outcome assessors from knowledge of which intervention a participant received and any information relating to whether the intended blinding was effective. Review authors' judgement: what is the risk of detection bias due to knowledge of the allocated interventions by outcome assessors?

Incomplete outcome data

Description: assess the completeness of outcome data for each main outcome, including attrition and exclusions from the analysis.

Review authors' judgement: what is the risk of attrition bias due to amount, nature, or handling of incomplete outcome data?

Selective reporting

Description: attempts were made to assess the possibility of selective outcome reporting by investigators.

Review authors' judgement: what is the risk of reporting bias due to selective outcome reporting?

Other sources of bias

We attempted to address sources of bias in other domains not covered by the tool. These included source of funding, conflicts of interest, and validity of outcome measures.

Review authors' judgement: what is the risk of bias due to problems not covered elsewhere in the table?

3.5.3.4 Measures of treatment effect

Dichotomous outcome data

We calculated risk ratio (RR) and 95% confidence intervals for dichotomous outcomes.

Continuous outcome data

For continuous outcomes, the Hedges' method was used to calculate standardized mean differences (SMDs) with individual study weights calculated as the inverse of the variance. To ensure that all scales were pointing in the same direction, we multiplied the mean value of one set by -1. As proposed by Cohen et al., an SMD of 0.2 represents a small effect, 0.5 a moderate effect and 0.8 a large effect (22).

3.5.3.5 Unit of analysis issues

Crossover trials

Since we calculated SMDs for all our continuous outcomes, we treated crossover studies as if they were parallel and computed a pooled SD. Although this method does not account for the correlation in crossover studies, it prevented any overestimation of effect sizes, which is desirable when computing SMDs. Carryover was not reported in any of the crossover studies.

Studies with multiple comparisons

For studies with more than two independent comparisons, such as, amphetamine versus placebo versus psychotherapy, we excluded the psychotherapy arm. We handled studies with multiple and correlated interventions, for example, or 10 mg dexamphetamine versus 20 mg of dexamphetamine versus placebo in the following way. For continuous outcomes of parallel studies, the means were calculated using the formulae described in Table 7.7.a of the *Cochrane Handbook for Systematic Reviews of Interventions* (23). For dichotomous outcomes of parallel studies, the number of events was added up across intervention arms. For continuous outcomes of crossover studies, both the means and the

standard deviations of the relevant intervention arms were averaged across the groups. For dichotomous outcomes of crossover studies, we randomly dropped one arm and used the other in the meta-analysis.

Studies with multiple time-points

We analysed studies separately according to their time frame. Time frames were denoted as 'short-term' (up to 6 months), medium term (between 6 and 12 months), and long-term (over 12 months). All but one study, Gillberg 2011 (25), were considered short-term. Since Gillberg 2011 was the only one that was considered medium-term, it was excluded from the meta-analysis.

3.5.3.6 Dealing with missing data

We e-mailed study authors up to 3 times (with at least 1 month between contacts) to obtain missing data. For those studies that did not report outcomes using intention-to-treat analysis and the missing data were unobtainable, meta-analysis was conducted as an available-case analysis, whereby data on only those participants whose results are known were included. For continuous outcomes, we used the sample size used to calculate the mean and standard deviations in the study. For studies that did not report SDs, it was calculated from p-values, confidence intervals or standard errors (as described in section 7.7.3.3 of the *Cochrane Handbook for Systematic Reviews of Interventions* (23). We did not use any imputations to deal with missing data.

3.5.3.7 Assessment of heterogeneity

Statistical heterogeneity was assessed by examining the I^2 index, which is used to quantify the degree of heterogeneity in a meta-analysis. We conducted a series of

subgroup analyses, which were selected *a priori* and based on preliminary evidence from other studies. These subgroups included: i) type of amphetamine formulation used; ii) amphetamine-release formulation; and iii) study funding source. We performed subgroup analysis when there were a sufficient number of studies, regardless of the degree of statistical heterogeneity present in the main analysis. We calculated a pooled effect size for each subgroup.

3.5.3.8 Assessment of reporting biases

Since none of the meta-analyses included a sufficient number of studies (i.e. ≥10) we did not explore the possibility of publication bias, the relationship between trial size and effect size, or chance using funnel plots.

3.5.3.9 Data synthesis

We synthesized the results in a meta-analysis using the random-effects model taking into account both within and between study variance. The inverse-variance method was used for continuous outcomes, and the Mantel-Haenszel method was used for dichotomous outcomes.

3.5.3.10 Subgroup analysis and investigation of heterogeneity

The following subgroup analyses were planned if there were a sufficient number of studies to warrant them:

1. Comorbidities: presence of comorbid ODD/CD or not. We did not conduct subgroup analysis for presence of comorbidities since more than half the studies failed to report whether or not they included or excluded participants with a comorbid condition.

2. Type of ADHD subtype: inattentive type, hyperactive-impulsive type, or combined type. We did not carry out this meta-analysis since the majority of studies failed to report this or did not subgroup their results according to this characteristic.
3. Type of amphetamine: dextroamphetamine, lisdexamphetamine, or mixed amphetamine salts. This analysis was conducted.
4. Type of drug release formulation: long-acting (extended release) or short-acting (immediate release). This analysis was conducted.
5. Funding source: with or without pharmaceutical industry funding. Since some studies failed to report their funding source, we grouped studies as 'industry funded', 'publicly funded', or 'not reported'. This analysis was conducted.

We conducted subgroup analyses on the following outcomes: total score on core symptom ADHD scale-Parent ratings; proportion of responders according to CGI-I; academic performance; retention: proportion of participants who completed the trial; proportion of participants who dropped out/withdrew due to an adverse event; proportion of participants experiencing decreased appetite; proportion of participants experiencing insomnia; proportion of participants experiencing abdominal pain; and proportion of participants experiencing headaches. We were unable to subgroup if all of the studies in a particular meta-analysis belonged to only one strata of any subgroup.

3.5.3.11 Sensitivity analysis

We planned to conduct the following sensitivity analyses:

1. Risk of bias assessment: each outcome meta-analysis was restricted to those studies with a low risk of bias. A study was defined as having a low risk of bias if all

domains of the risk of bias tool scored a low risk of bias. Since no studies met this criterion, we did not carry out this sensitivity analysis.

2. Unpublished versus published studies. Since no unpublished studies were included in this review, we did not perform this sensitivity analysis.
3. Imputed SD versus a lower and higher imputed SD (in the event of missing data). Since we did not have to impute any SDs in this review, we did not conduct this sensitivity analysis.

3.6 RESULTS

3.6.1 Results of the search

Figure 3.12.1 summarizes the flow of studies through the screening process. The results of the electronic databases retrieved 5673 publications, while the search through other sources yielded 191 studies. After removing duplicates, 4228 references were identified for further consideration. After screening the titles and available abstracts, the full texts for 293 studies were examined. Of those 293 studies, 20 studies were included (24-27; 32-47), which accounted for 24 publications, since one study had 4 publications associated with it (26) and another had 2 publications associated with it (27) (one of which was the pilot, and the other was the full study). Two clinical trials were ongoing when we completed this review (28, 29), and although recruitment ended for both of these trials, none of the results have been published. Both authors were contacted 3 times yielding no response. One non-English study is awaiting classification since information on whether it was randomized and whether participants had a formal diagnosis of ADHD is needed (30). This information was unattainable given the inability to contact the author. Another study also awaits classification as only the abstract has been published

(31). Information on whether treatments were randomized and whether participants had a formal diagnosis is needed. Authors were contacted three times yielding no response.

3.6.2 Description of included studies

Twenty studies met the inclusion criteria. Seven studies were parallel-group trials, while 13 studies were crossover. Nine studies included a single comparison of an amphetamine derivative versus placebo, 2 studies compared more than one amphetamine derivative with placebo, and 10 studies compared more than one dose of an amphetamine derivative with placebo. Seventeen studies were included in the meta-analysis. See Table 3.11.1 for a more detailed description of included studies.

3.6.2.1 Patients

A total of 2183 participants were randomized to relevant interventions, with 1569 (72%) of them being male. The age of the participants ranged from 3-17 years.

3.6.2.2 Interventions

Eleven studies assessed mixed amphetamine salts, 6 studies used dextroamphetamine and 5 studies looked at lisdexamphetamine (2 studies assessed 2 amphetamine derivatives). Ten studies randomized participants to set doses or dosing schedules, 7 studies used weight-based dosing, while 4 studies titrated participants to their optimal dose (1 study used both weight-based and titration).

3.6.2.3 Duration

Study intervention length ranged from 14 to 365 days, with a median of 28 days. Only one study was longer than 63 days.

3.6.2.4 Setting

Eight studies were multi-center trials and nineteen were conducted in the United States.

3.6.3 Excluded studies

We excluded a total of 270 studies (see Table 3.11.2 for reasons for exclusion for selected studies).

3.6.4 Risk of bias in included studies

A more in-depth risk of bias assessment for each study can be found in the Table 3.11.1. In addition, Figure 3.12.2 provides a summary of this assessment.

Random sequence generation

Only two studies reported on how the random sequence was generated, and were assessed as having a 'low' risk of bias in this domain. The other 18 studies were rated as 'unclear', as they did not adequately describe their methods of randomization.

Allocation concealment

Three studies described the methods used to conceal the allocation sequence, and were rated as having a 'low' risk of bias in this domain. The rest of the studies were assessed as 'unclear', as they did not sufficiently describe their methods of allocation concealment.

Performance bias

Although blinding was intended in all of the studies, we assessed risk of bias by how authors described their amphetamine and placebo capsules. Eight studies were rated as 'low' in this domain. The other 12 studies were marked as being unclear since they were not explicit about the similarities between the two interventions.

Detection bias

Only 2 studies explicitly stated that outcome assessors were blind to interventions, and were therefore marked as having a 'low' risk of bias. The other 18 studies were rated as 'unclear' since they were not explicit about which parties were blind to the intervention assignment.

Incomplete outcome data

In this domain, 12 studies adequately addressed their drop-outs and statistical methods used to compensate for the drop-outs, and were rated as having a 'low' risk of bias. Five studies failed to provide reasons for their drop-outs and failed to address any exclusions from their analyses, and were therefore rated as having a 'high' risk of bias. The other 3 studies did not discuss drop-outs in their reports and were rated as 'unclear'.

Selective reporting

Fifteen studies were assessed as having a 'low' risk of bias in this domain as they reported on all outcomes discussed in the methods and provided adequate data for each of the outcomes (i.e. measure of effect and variance). Five studies were assessed as 'high' since they failed to report on all the outcomes mentioned in their methods, and provide adequate data.

Other potential sources of bias

Twelve studies were rated as having a 'high' risk of bias in this domain since they reported that they were funded by and/or affiliated with the pharmaceutical industry. Two studies were rated as 'unclear' since the validity of their primary outcomes was not described. The other 6 studies appeared to be free of other potential sources of bias.

3.6.5 Effects of interventions

3.6.5.1 Primary outcome

Given that the primary outcome is change in core ADHD symptoms (inattention, hyperactivity, impulsivity), as measured by a validated scale rated by children, parents, teachers, clinicians, or assessors, a series of meta-analyses were conducted (Figures 3.12.3.1 to 3.12.3.12 [Analyses 1.1 to 1.12]). In all 12 outcomes, amphetamines were significantly superior to placebo. These outcomes include: total ADHD symptom score-parent ratings (Analysis 1.1; SMD -0.40 [95%CI -0.53 to -0.27]); parent ratings of hyperactivity/impulsivity (Analysis 1.2; SMD -0.63 [95% CI-0.84 to -0.41]); parent ratings of inattention (Analysis 1.3; SMD -1.02 [95%CI -1.30 to -0.74]); total ADHD symptom score-teacher ratings (Analysis 1.4; SMD -0.55 [95%CI -0.83 to -0.27]); teacher ratings of hyperactivity/impulsivity (Analysis 1.5; SMD -1.13 [95%CI -1.63 to -0.62]); teacher ratings of inattention (Analysis 1.6; SMD -1.43 [95%CI -2.35 to -0.52]); total ADHD symptom score-clinician ratings (Analysis 1.7; SMD -0.84 [95%CI -1.32 to -0.36]); clinician ratings of hyperactivity/impulsivity (Analysis 1.8; SMD -0.75 [95%CI -1.28 to -0.23]); clinician ratings of inattention (Analysis 1.9; SMD -0.78 [95%CI -1.26 to -0.30]); total ADHD symptom score-investigator ratings (Analysis 1.10; SMD -0.49 [95%CI -0.76 to -0.23]); investigator ratings of hyperactivity/impulsivity (Analysis 1.11; SMD -1.21 [95%CI -1.72 to -0.69]); and investigator ratings of inattention (Analysis 1.12; SMD -0.41 [95%CI -0.63 to -0.19]). It is important to note that the majority of these meta-analyses included between 1-3 studies, and that Analyses 1.7, 1.8 and 1.9 had considerable heterogeneity present at an I^2 of 88%, 90% and 88% respectively. Only 2 outcomes, total ADHD symptom score-parent ratings (Analysis 1.1; Figure 3.12.3.1) and

total ADHD symptom score-teacher ratings (Analysis 1.4; Figure 3.12.3.4) included more than 3 studies (6 and 5 respectively). Statistical heterogeneity was non-significant in each of these analyses.

3.6.5.2 Secondary outcomes

We conducted 5 meta-analyses that explored secondary outcomes (Figures 3.12.3.13 to 3.12.3.17 [Analyses 1.13 to 1.17]. Out of these, 3 outcomes yielded statistically significant results all in favour of amphetamine. These outcomes included: clinical global impression severity score (CGI-S) (Analysis 1.13; SMD -0.86 ;95%CI -1.72 to -0.01]); proportion of responders (Analysis 1.14; RR 3.12 [95%CI 2.32 to 4.20]); and academic performance (Analysis 1.16; SMD 0.51 [95%CI 0.31 to 0.70]). The other two outcomes, quality of life (Analysis 1.15; SMD -0.01 [95%CI -0.27 to 0.25]) and retention in the trial (Analysis 1.17; RR 1.05 [95%CI 0.99 to 1.11]) did not show significant results.

3.6.5.3 Adverse events

Although the proportion of participants who experienced at least one adverse event was higher in the amphetamine groups as compared to placebo (Analysis 1.18, Figure 3.12.3.18; RR 1.31 [95%CI 1.16 to 1.48]), there was no difference in the proportion of participants who withdrew due to an adverse event between the amphetamine and placebo groups (Analysis 1.19; RR 1.74 [95%CI 0.86 to 3.52]). We meta-analysed the most commonly reported adverse events (Analysis 1.20 to Analysis 1.25). The proportion of participants who experienced decreased appetite (Analysis 1.20; RR 7.44 [95%CI 2.99 to 18.48]), insomnia/trouble sleeping Analysis 1.21; RR 3.67 [95%CI 1.83 to 7.38]), and abdominal pain (Analysis 1.22; RR 1.65 [95%CI 1.17 to 2.31]) was significantly higher

in the amphetamine groups as compared to placebo. There was no difference in the proportion of participants who experienced headaches (Analysis 1.23; RR 0.97 [95%CI 0.75 to 1.24]), anxiety/nervousness (Analysis 1.24; RR 1.35 [95%CI 0.67 to 2.73]), and nausea (Analysis 1.25; RR 1.38 [95%CI 0.77 to 2.48]) between amphetamine and placebo groups.

3.6.6 Subgroup analyses

We conducted a series of subgroup analyses according to type of amphetamine (dexamphetamine, lisdexamphetamine, and mixed amphetamine salts) (Figures 3.12.4.1 to 3.12.4.12 [Subgroup Analyses 2.1 to 2.12). We found no significant between-group differences in any of the outcomes. It is important to note that although more participants in the lisdexamphetamine group experienced headaches in comparison to placebo, fewer participants experienced headaches in the mixed amphetamine salts group (Subgroup Analysis 2.12, Figure 3.12.4.12), though this difference was not statistically significant (P=0.11).

The influence of amphetamine release formulation was explored by subgrouping by long-acting versus short-acting formulations (Figures 3.12.5.1-3.12.5.6 [Subgroup Analyses 3.1 to 3.6]). There were no significant between-group differences in any of the outcomes. Long-acting formulation was associated with a slightly higher retention as compared to the short-acting formulation, however, this was not statistically significant (p=0.22) (Subgroup Analysis 3.4).

We wanted to explore the influence of study funding source by subgrouping studies according to those that were industry funded versus publicly funded (Figures 3.12.6.1-3.12.6.3 [Subgroup Analyses 4.1 to 4.3]). Since 5 studies did not report their source of funding, another subgroup was introduced as 'not reported'. Although no significant between-group differences were found, it is important to note that out of the 15 studies that did report their source of funding, 12 studies were funded by industry thereby causing an imbalance in these subgroups.

3.6.7 Sensitivity analysis

As described in section 3.5.3.11, no sensitivity analyses were carried out.

3.7 DISCUSSION

3.7.1 Summary of main results

Twenty studies were included in this review, and 17 in the meta-analysis. Overall, this review found that amphetamines are an effective short-term treatment option for the core symptoms of ADHD. The largest effects observed (i.e. an SMD >0.7) were parent ratings of inattention; teacher ratings of inattention and hyperactivity/impulsivity; clinician ratings of total ADHD symptoms, inattention, and hyperactivity/impulsivity; and investigator rating of hyperactivity/impulsivity. These results must be interpreted with caution, as most of these meta-analyses included very few studies. The largest meta-analyses included parent and teacher ratings of total ADHD symptoms, both of which yielded low to moderate effects. Furthermore, there was a lot of variation in the amphetamine derivatives and release formulations utilized in the included studies. As such, we conducted subgroup analyses to assess their differences. Minimal between-

group differences were found, however, it is important to note that most studies assessed mixed amphetamine salts and long-term release formulations lending to an imbalance in the subgroups. There was no evidence of a beneficial effect of amphetamines on quality of life; however, only one included study measured this.

The most commonly reported adverse events in the primary studies were meta-analysed. These included decreased appetite, insomnia, abdominal pain, headaches, anxiety, and nausea/vomiting. Meta-analysis revealed that most adverse events occurred significantly more often in the amphetamine groups than in the placebo groups.

3.7.2 Overall completeness and applicability of evidence

This review only focused on the amphetamine versus placebo comparison. While it is important to assess amphetamines versus other active therapies such as other stimulants, psychotherapy or anti-depressants, we believe it was important to first establish whether or not amphetamines are superior to placebo. We were unable to assess the long-term efficacy of amphetamines (i.e. beyond 12 months of use). The median duration of included studies was 4 weeks long (25). Short-term trials are particularly problematic for chronic conditions such as ADHD, as children will likely be on stimulant medications for much longer periods than what has been studied.

As mentioned earlier, adverse events occurred more frequently when participants were on amphetamines versus when they were on placebo, however, these results must be cautiously interpreted given the poor reporting around adverse events in the primary studies. Some studies only reported on adverse events that were experienced by a certain percentage of participants, thereby potentially ignoring additional adverse events

experienced at less than that fraction. Furthermore, many studies were unclear regarding their methods of collecting adverse events and whether they assessed the causality of these adverse events as it related to the interventions. Heterogeneity of terms used to describe adverse events was also a major hurdle when conducting this review, and limited our ability to appropriately synthesize the data. Finally, as with efficacy data, we were unable to assess the long-term safety profile of amphetamines given the lack of long-term trials.

The external validity of our results was also limited by the eligibility criteria of the included studies. Since we excluded studies that included participants with comorbidities other than CD, ODD, anxiety, and depression we cannot extrapolate the results of our review to patients with other commonly occurring comorbidities such as depression, and tic disorder. Generalizability of the results is also compromised given the characteristics of included studies. In those studies where ADHD subtype was reported, 79% of included participants were of the combined subtype, therefore extrapolation of results to children who are predominantly inattentive or hyperactive/impulsive is questionable. Moreover, since 72% of study participants were male, this further limits the generalizability of the results to females.

Since 50% of included studies failed to report on the ADHD subtype make-up of their included participants, we were unable to make any conclusions regarding the potential heterogeneity of effect of different formulations of amphetamines across the different ADHD subtypes. Furthermore, since primary studies did not subgroup their results

according to important prognostic factors such as age and gender, we were unable to subgroup our meta-analyses, which are particularly relevant for clinicians.

3.7.3 Quality of the evidence

Although our results favour the use of amphetamines, we recognize that the strength of our results depends on the internal validity of the included RCTs in our review. None of the included studies scored a low risk of bias on all domains of the Cochrane Risk of Bias tool, which can result in an overestimation of treatment effect (48). Moreover, most studies failed to report on how the random sequence was generated (90%), how allocation was concealed (85%), methods used to blind participants, personnel (55%) and outcome assessors (90%). As such, we were unable to determine whether it was a reporting problem versus a study design problem.

In addition, the results of this review may have been influenced by the fact that 60% of the included studies were industry funded. This has been strongly associated with an overestimation of treatment effect in favour of the sponsor's interest, further potentially distorting the true effect of amphetamines (49).

3.7.4 Potential biases in the review process

Limitations of our review include not being able to account for correlation in crossover studies given the formula for calculating SMDs, thereby resulting in a more conservative treatment effect for crossover trials.

Given the few number of studies included in the meta-analyses, we were unable to assess reporting bias using funnel plots, and were therefore unable to assess whether publication

status played a role in this review, which may have led to an overestimate of treatment effects. Furthermore, the exclusion of one potentially eligible non-English study may have also biased our findings. Egger et al (50) found that non-English studies tend to be negative; therefore, excluding them may have yielded an overestimation of treatment effects. On the other hand, other researchers have found that excluding trials reported in languages other than English do not significantly affect the results of a meta-analysis (51).

3.7.5 Agreements and disagreements with other studies or reviews

Two previously conducted systematic reviews were identified prior to conducting ours (17, 18). Charach et al. assessed only the long-term (i.e. >12 months) efficacy and safety of amphetamines for pediatric ADHD. Furthermore, they included all study designs (observational studies, open-label extensions, and RCTs) in their review, of which only one was an RCT, which was also included in this systematic review (24). Miller et al. also systematically assessed amphetamines for pediatric ADHD, however, reviewers only included studies that assessed the dexamphetamine derivative of amphetamines. Furthermore, this review was published in 1999, making it over 14 years old. As such, Miller et al included only one relevant RCT in their review, which was also included in this review (26).

3.8 CONCLUSION

3.8.1 Implications for practice

Although this review demonstrates that the effect of amphetamines in improving core symptoms of ADHD in the short term is generally favourable, clinicians must weigh the

benefits of this intervention with its safety profile. In addition, clinicians must consider the heterogeneity of treatment response across their patients depending on age, sex, presence of comorbidities, and ADHD subtype. Broad generalizations regarding the efficacy of amphetamines should be avoided. Furthermore, this review does not provide evidence that supports any one amphetamine derivative over another, and does not reveal any differences between long-acting and short-acting formulations. Given that long-acting formulations can cost up to 15 times more, further research is needed to investigate their cost-benefits ratio, in particular whether they do in fact achieve promise of greater compliance.

3.8.2 Implications for research

Future RCTs should be longer in duration in order to explore the long-term safety and efficacy of amphetamines. In addition, future studies should not only focus on symptom management, but also global outcomes such as quality of life, academic performance, and persistence of ADHD symptoms into adulthood. It would also be beneficial for future studies to subgroup their results based on important prognostic factors such as age and gender. Researchers should consult the Consolidated Standards of Reporting Trials (CONSORT) statement when designing their study and reporting their methods and results so that an appropriate risk of bias assessment can be made allowing for a more robust interpretation (52).

3.9 ACKNOWLEDGEMENTS

The authors wish to thank Margaret Anderson and Soleil Surrette for their assistance in developing and reviewing the search strategy.

This review was produced within the Cochrane Developmental, Psychosocial and Learning Problems Group.

3.10 REFERENCES

1. Polanczyk G, Silva de Lima M, Horta BL, Biederman J, Rohde LA. The worldwide prevalence of ADHD: a systematic review and metagression analysis. *American Journal of Psychiatry* 2007;164(6):942-8.

2. American Psychiatric Association. Diagnostic and Statistical Manual of Mental Disorders (DSM-IV-TR). 4th text revision edition. Washington, DC: American Psychiatric Association, 2000.

3. American Academy of Pediatrics, Subcommittee on Attention-Deficit/Hyperactivity Disorder, Steering Committee on Quality Improvement and Management. ADHD: Clinical Practice Guideline for the Diagnosis, Evaluation, and Treatment of Attention-Deficit/Hyperactivity Disorder in Children and Adolescents. *Pediatrics* 2011;128:1007-22.

4. Agency for Health Care Policy and Research. Diagnosis of attention-deficit/hyperactivity disorder. Summary: technical review: number 3. www.ahrq.gov/clinic/epcsums/adhdsutr.htm (accessed 1 May 2012).

5. Mayes SD, Calhoun SL, Bixler EO, Vgontzas AN, Mahr F, Hillwig-Garcia J, et al. ADHD subtypes and comorbid anxiety, depression and oppositional-defiant disorder: differences in sleep problems. *Journal of Pediatric Psychology* 2009;34(33):328-37.

6. National Institute of Health. National Institutes of Health consensus development conference statement: diagnosis and treatment of attention-deficit/hyperactivity disorder (ADHD). *Journal of the American Academy of Child and Adolescent Psychiatry* 2000;39(2):182-97.

7. Edwards L, Salant V, Howard VF, Brougher J, McLaughlin TF. Effectiveness of self-management on attentional behavior and reading comprehension for children with attention deficit disorder. *Child and Family Behavior Therapy* 1995;17(2):1-17.

8. Faraone SV, Biederman J, Mick E. The age-dependent decline of attention-deficit/hyperactivity disorder: a meta-analysis of follow-up studies. *Psychological Medicine* 2006;36:159–165.

9. Wigal T, Swanson JM, Regino R, Lerner MA, Soliman I, Steinoff K, et al. Stimulant medications for the treatment of ADHD: efficacy and limitations. *Mental Retardation and Development Disabilities Research Reviews* 1999;5(3):215-24.

10. Buck ML. Amphetamines in the treatment of attention-deficit/hyperactivity disorder. *Pediatric Pharmacotherapy* 2002;8(3):1-3.

11. The Medical Letter Online. Lisdexamfetamine dimesylate (Vyvanase) for ADHD. *The Medical Letter on Drugs and Therapeutics* 2007;49(1265):58-9.

12. Swanson JM, Kinsbourne M, Nigg J, et al. Etiologic subtypes of attention-deficit/hyperactivity disorder: brain imaging, molecular genetic and environmental factors and the dopamine hypothesis. *Neuropsychology Reviews* 2007;17(1):39-59

13. Anderson SW, Bechara A, Damasio H, Tranel D, Damasio AR. Impairment of social and moral behavior related to early damage in human prefrontal cortex. *Natural Neuroscience* 1999;2(11):1032-7.

14. Volkow ND, Wang GJ, Newcorn J, et al. Depressed dopamine activity in caudate and preliminary evidence of limbic involvement in adults with attention-deficit/hyperactivity disorder. *JAMA Psychiatry* 2007;64(8):932-940.

15. Arnsten AFT. Stimulants: therapeutic actions in ADHD. *Neuropsychopharmacology* 2006;31(11):2376-83.

16. Castells X, Ramos-Quiroga JA, Bosch R, Nogueira M, Casas M. Amphetamines for Attention Deficit Hyperactivity Disorder (ADHD) in adults. *Cochrane Database of Systematic Reviews* 2011, Issue 6. Art. No.: CD007813. DOI: 10.1002/14651858.CD007813.pub2.

17. Charach A, Dashti B, Carson P, Booker L, Lim CG, Lillie E, et al. Attention deficit hyperactivity disorder: effectiveness of treatment in at-risk preschoolers; long-term effectiveness in all ages; and variability in prevalence, diagnosis, and treatment.

Agency for Healthcare Research and Quality 2011.

18. Miller A, Lee S, Raina P, Klassen A, Zupancic J, Olsen L. A review of therapies for attention-deficit/hyperactivity disorder. www.cadth.ca/en/products/health-technology-assessment/publication/20 (accessed 11 February 2014).

19. American Psychiatric Association. Diagnostic and Statistical Manual of Mental Disorders (DSM-III-R). 3rd edition. Washington, DC: American Psychiatric Association, 1987.

20. American Psychiatric Association. Diagnostic and Statistical Manual of Mental Disorders (DSM-IV-TR). 4th text revision edition. Washington, DC: American Psychiatric Association, 2000.

21. Review Manager (RevMan) [Computer program]. Version 5.1. Copenhagen: The Nordic Cochrane Centre, The Cochrane Collaboration, 2011.

22. Cohen J. Statistical Power Analysis in Behavioral Sciences (2nd edition. Hillsdale (NJ): Lawrence Erlbaum Associates Inc., 1988.

23. Higgins JPT, Green S (editors). Cochrane Handbook for Systematic Reviews of Interventions. Chichester: John Wiley & Sons, 2011.

24. Gillberg C, Melander H, Knorring AL, Janols LO, Thernlund G, Hagglof B, et al. Long-term stimulant treatment of children with attention-deficit hyperactivity disorder symptoms. A randomized, double-blind, placebo-controlled trial. *Archives of General Psychiatry* 1997;54(9):857-64.

25. Biederman J, Lopez FA, Boellner SW, Chandler MC. A randomized, double-blind, placebo-controlled, parallel-group study of SLI381 (Adderall XR) in children with attention-deficit/hyperactivity disorder. *Pediatrics* 2002;110(2):258-66.

26. Borcherding BG, Keysor CS, Rapoport JL, Elia J, Amass J. Motor/vocal tics and compulsive behaviors on stimulant drugs: is there a common vulnerability? *Psychiatry Research* 1990;33(1):83-94.

Elia J, Borcherding BG, Potter WZ, Mefford IN, Rapoport JL, Keysor CS. Stimulant drug treatment of hyperactivity: biochemical correlates. *Clinical Pharmacology and Therapeutics* 1990;48(1):57-66.

Elia J, Borcherding BG, Rapoport JL, Keysor CS. Methylphenidate and dextroamphetamine treatments of hyperactivity: are there true nonresponders? *Psychiatry Research* 1991;36(2):141-55.

Elia J, Welsh PA, Gullotta CS, Rapoport JL. Classroom academic performance: improvement with both methylphenidate and dextroamphetamine in ADHD boys. *Journal of Child Psychology and Psychiatry* 1993;34(5):785-804.

27. Donnelly M, Rapoport JL, Ismond DR. Fenfluramine treatment of childhood attention deficit disorder with hyperactivity: a preliminary report. *Psychopharmacology Bulletin* 1986;22(1):152-4.

Donnelly M, Rapoport JL, Potter WZ, Oliver J, Keysor CS, Murphy DL. Fenfluramine and dextroamphetamine treatment of childhood hyperactivity. Clinical and biochemical findings. *Archives of General Psychiatry* 1989;46(3):205-12.

28. Fanton J, Waslick B, Harvey E. The 49th Annual National Institute of Mental Health (NIMH) New Clinical Drug Evaluation Unit (NCDEU) meeting Hollywood, Florida, June29-July 2, 2009: posters most relevant to child and adolescent psychopharmacology. *Journal of Child and Adolescent Psychopharmacology* 2009;19(6):786.

29. Waxmonsky J. A randomized, double-blind, placebo-controlled, crossover, laboratory classroom study to evaluate the safety and efficacy of d-amphetamine transdermal drug delivery system (d-ATS) compared to placebo in children and adolescents with ADHD. http://clinicaltrials.gov/ct2/show/record/NCT01711021 (accessed 11 February 2014).

30. Glos J. [Amphetamines in the treatment of hyperactive and instable children]. *Ceskoslovenska pediatrie* 1973;28(10):559-60.

31. Itil TM, Simeon J. Proceedings: Computerized EEG in the prediction of outcome of drug treatment in hyperactive childhood behavior disorders. *Psychopharmacology bulletin* 1974;10(4):36.

32. Shekim WO, Bylund DB, Alexson J, Glaser RD, Jones SB, Hodges K, et al. Platelet MAO and measures of attention and impulsivity in boys with attention deficit disorder and hyperactivity. *Psychiatry Research* 1986;18(2):179-88.

33. Biederman J, Krishnan S, Zhang Y, McGough JJ, Findling RL. Efficacy and tolerability of lisdexamfetamine dimesylate (NRP-104) in children with attention-deficit/hyperactivity disorder: a phase III, multicenter, randomized, double-blind, forced-dose, parallel-group study. *Clinical Therapeutics* 2007;29(3):450-63.

34. Findling RL, Childress AC, Cutler AJ, Gasior M, Hamdani M, Ferreira-Cornwell MC, et al. Efficacy and safety of lisdexamfetamine dimesylate in adolescents with attention-deficit/hyperactivity disorder. *Journal of the American Academy of Child and Adolescent Psychiatry* 2011;50(4):395-405.

35. Biederman J, Boellner SW, Childress A, Lopez FA, Krishnan S, Zhang Y. Lisdexamfetamine dimesylate and mixed amphetamine salts extended-release in children with ADHD: a double-blind, placebo-controlled, crossover analog classroom study. *Biological Psychiatry* 2007;62(9):970-6.

36. Manos MJ, Short EJ, Findling RL. Differential effectiveness of methylphenidate and Adderall in school-age youths with attention-deficit/hyperactivity disorder. *Journal of the American Academy of Child and Adolescent Psychiatry* 1999;38(7):813-9.

37. James RS, Sharp WS, Bastain TM, Lee PP, Walter JM, Czarnolewski M, et al. Double-blind, placebo-controlled study of single-dose amphetamine formulations in ADHD. *Journal of the American Academy of Child and Adolescent Psychiatry* 2001;40(11):1268-76.

38. Nemzer ED, Arnold E, Votolato NA, McConnell H. Amino acid supplementation as therapy for attention deficit disorder. *Journal of the American Academy of Child Psychiatry* 1986;25(4):509-13.

39. Pliszka SR, Browne RG, Olvera RL, Wynne SK. A double-blind, placebo-controlled study of Adderall and methylphenidate in the treatment of attention-deficit/hyperactivity disorder. *Journal of the American Academy of Child and Adolescent Psychiatry* 2000;39(5):619-26.

40. Sharp WS, Walter JM, Marsh WL, Ritchie GF, Hamburger SD, Castellanos FX. ADHD in girls: clinical comparability of a research sample. *Journal of the American Academy of Child and Adolescent Psychiatry* 1999;38(1):40-7.

41. Short EJ, Manos MJ, Findling RL, Schubel EA. A prospective study of stimulant response in preschool children: insights from ROC analyses. *Journal of the American Academy of Child and Adolescent Psychiatry* 2004;43(3):251-9.

42. Swanson JM, Wigal S, Greenhill LL, Browne R, Waslik B, Lerner M, et al. Analog classroom assessment of Adderall (R) in children with ADHD. *Journal of the American Academy of Child and Adolescent Psychiatry* 1998;37(5):519-26.

43. Wigal SB, Kollins SH, Childress AC, Squires L, Brams M, Childress A, et al. A 13-hour laboratory school study of lisdexamfetamine dimesylate in school-aged children with attention-deficit/hyperactivity disorder. *Child and Adolescent Psychiatry and Mental Health* 2009;3(1):17.

44. Barkley RA, Connor DF, Kwasnik D. Challenges to determining adolescent medication response in an outpatient clinical setting: comparing Adderall and methylphenidate for ADHD. *Journal of Attention Disorders* 2000;4(2):102-13.

45. Giblin GM, Strobel AL. Effect of lisdexamphetamine dimesylate on sleep in children with ADHD. *Journal of Attention Disorders* 2011;15(6):491-8.

46. McCracken JT, Biederman J, Greenhill LL, Swanson JM, McGough JJ, Spencer TJ, et al. Analog classroom assessment of a once-daily mixed amphetamine formulation, SLI381 (Adderall XR), in children with ADHD. *Journal of the American Academy of*

Child and Adolescent Psychiatry 2003;42(6):673-83.

47. Spencer TJ, Wilens TE, Biederman J, Weisler RH, Read SC, Pratt R. Efficacy and safety of mixed amphetamine salts extended release (Adderall XR) in the management of attention-deficit/hyperactivity disorder in adolescent patients: a 4-week, randomized, double-blind, placebo-controlled, parallel-group study. *Clinical Therapeutics* 2006;28(2):266-79.

48. Hartling L, Ospina M, Liang Y, Dryden DM, Hooton N, Krebs Seida J, et al. Risk of bias versus quality assessment of randomized controlled trials: cross sectional study. *British Medical Journal* 2009;339:1-6.

49. Lundh A, Sismondo S, Lexchin J, Busuioc OA, Bero L. Industry sponsorship and research outcome. *Cochrane Database of Systematic Reviews* 2012(12):Art. No.: MR000033. DOI: 10.1002/14651858.MR000033.pub2.

50. Egger M, Smith GD, Schneider M, Junker C, Lengeler C, Antes G. Language bias in randomized controlled trials published in English and German. *The Lancet* 1997;350:326-9.

51. Moher D, Pham B, Lawson ML, Klassen TP. The inclusion of reports of randomized trials published in languages other than English in systematic reviews. *Health*

Technology Assessment 2003;7:1-90.

52. Schulz KF, Altman DG, Moher D for the CONSORT Group. CONSORT 2010 statement: updated guidelines for reporting parallel group randomized trials. *British Medical Journal* 2010;340:c332.

3.11 TABLES

Table 3.11.1 Characteristics of included studies

1. Barkley 2000		
Methods	Double-blind, placebo-controlled, crossover, randomized clinical trial Country: United States Statistical methods: per protocol	
Participants	N=46* participants with an ADHD diagnosis according to the DSM-IV criteria Psychiatric comorbid conditions: NR Age range: 12-17 Mean age (SD): 14 (NR) Male: 30 (86%) ADHD subtype: NR	
Interventions	Five interventions: 1. Mixed amphetamine salts (short-acting), 5 mg bid 2. Mixed amphetamine salts (short-acting), 10 mg bid 3. Methylphenidate, 5 mg bid 4. Methylphenidate, 10 mg bid 5. Placebo Duration: 35 days (five 7-day treatment periods)	
Outcomes	Relevant outcomes: ADHD core symptom severity assessed with ADHD-RS (parents and teachers) Clinical impression assessed with CGI-Improvement Adverse events Other outcomes: ODD symptom severity assessed with an ODD rating scale (not specified) Continuous performance test Response inhibition and interference control assessed with: Stroop Word-Color Association Test	
Other	Authors' affiliation: University Study funding: Pharmaceutical industry *Clinical characteristics were only reported on participants who completed the trial (n=35)	
Risk of bias assessment		
Bias	*Authors' judgement*	*Support for judgement*
Random sequence generation (selection bias)	Unclear	Method of sequence generation is not described
Allocation concealment (selection bias)	Unclear	Method of allocation concealment is not described
Incomplete outcome data	High	Only data for completers is included in

(attrition bias)		the analysis. For one of the primary outcomes, only 37% of randomized participants were included in the analysis.
Selective reporting (reporting bias)	Low	Data is provided on all outcomes listed in the methods. Study appears to be free of selective reporting.
Other bias	High	Study was funded by industry.
Blinding of participants and personnel (performance bias)	Unclear	Blinding of participants and personnel is not adequately described
Blinding of outcome assessment (detection bias)	Unclear	Blinding of outcome assessment is not described.
2. Biederman 2002		
Methods	Double-blind, multi-center, placebo-controlled, parallel-group, randomized clinical trial Country: United States Number of study sites: 47 Statistical methods: modified ITT (last observation carried forward)	
Participants	N=584* participants with an ADHD diagnosis according to DSM-IV criteria Psychiatric comorbid conditions: NR Age range: 6-12 years Mean age (SD):8.6 (1.7) Male: 434 (77%) ADHD subtype: Hyperactive-Impulsive: 26 (5%); Inattentive: 12 (2%); Combined: 523 (93%)	
Interventions	Four interventions: 1. Mixed amphetamine salts (long-acting), 10 mg/day, N=119 2. Mixed amphetamine salts (long-acting), 20 mg/day (10 mg/day for week 1 with forced dose escalation to 20 mg/day in weeks 2-3), N=105 3. Mixed amphetamine salts (long-acting), 30 mg/day (10 mg/day in week 1, 20 mg/day in week 2, 30 mg/day in week 3), N=112 4. Placebo, N=173 Duration: 21 days	
Outcomes	Relevant outcomes: ADHD core symptom severity assessed with: Conners' Global Index Scale-Teacher, Conners' Global Index Scale-Parent Clinician impression: CGI-Improvement Retention: proportion of participants who completed the trial Number of participants who experienced at least one adverse event Number of participants who dropped out due to any adverse event Adverse events Other outcomes:	

	Parent Global Assessment	
Other	Authors' affiliations: University Study funding: Pharmaceutical industry *clinical characteristics are only provided on individuals who were included in the primary efficacy analysis (n=563)	
Risk of bias assessment		
Bias	*Authors' judgement*	*Support for judgement*
Random sequence generation (selection bias)	Unclear	Method of sequence generation is not described
Allocation concealment (selection bias)	Unclear	Method of allocation concealment is not described
Incomplete outcome data (attrition bias)	Low	Attrition was moderate (13%). Reasons for attrition are reported and 96% of randomized participants were included in the primary analysis.
Selective reporting (reporting bias)	Low	Data is provided on all outcomes listed in the methods. Study appears to be free of selective reporting.
Other bias	High	Study was funded by industry.
Blinding of participants and personnel (performance bias)	Unclear	Blinding of participants and personnel is not adequately described
Blinding of outcome assessment (detection bias)	Unclear	Blinding of outcome assessment is not described.
3. Biederman 2007a		
Methods	Double-blind, multi-center, placebo-controlled, crossover, randomized clinical trial Country: United States Number of study sites: 4 Statistical methods: modified ITT, all randomized subjects who had at least one post-randomization score on primary outcome measure	
Participants	N=52 participants with an ADHD diagnosis according to the DSM-IV-TR criteria Psychiatric comorbid conditions: NR Age range: 6-12 years Mean age (SD): 9.1 (1.7) Male: 33 (64%) ADHD subtypes: Combined: 52 (100%)	
Interventions	Three interventions: 1. Mixed amphetamine salts (long-acting), either 10 mg/day, 20 mg/day or 30 mg/day qd (determined by dose optimization period)* 2. Lisdexamphetamine (long-acting), either 30 mg/day, 50 mg/day or 70 mg/day qd (determined by dose optimization period) 3. Placebo Duration: 21 days (three 7-day treatment periods)	

Outcomes	Relevant outcomes: ADHD core symptom severity assessed with: SKAMP-attention subscale Academic performance assessed with: Permanent Product Measure of Performance Clinical impression assessed with: CGI-severity and CGI-improvement scales Retention: number of participants who completed the study Adverse events Other outcomes: Conduct problems assessed with: SKAMP-deportment subscale Vital signs (blood pressure, pulse)
Other	ClinicalTrials.gov Identifier: NCT00557011 Authors' affiliation: university and pharmaceutical industry Study funding: Pharmaceutical industry *Mixed amphetamine salts-extended release was randomly chosen to represent the amphetamine group in this study for binary outcomes

Risk of bias assessment		
Bias	*Authors' judgement*	*Support for judgement*
Random sequence generation (selection bias)	Unclear	Method of sequence generation is not described
Allocation concealment (selection bias)	Low	Method of allocation concealment involved pre-packaged, serially-numbered drug kits, in which the next participant enrolled received the next available drug kit. Drug kits were prepared by a third party.
Incomplete outcome data (attrition bias)	Low	Although the primary only included trial completers, attrition was low at 4%. Reasons for drop-outs are reported.
Selective reporting (reporting bias)	Low	Data is provided on all outcomes listed in the protocol. Study appears to be free of selective reporting.
Other bias	High	Study was funded by industry and all authors are or have been affiliated with the pharmaceutical company that funded the study.
Blinding of participants and personnel (performance bias)	Unclear	Blinding of participants and personnel is not described
Blinding of outcome assessment (detection bias)	Unclear	Blinding of outcome assessment is not described.

4. Biederman 2007b	
Methods	Double-blind, multi-center, placebo-controlled, parallel-group

	randomized clinical trial Country: United States Number of study sites: 40 Statistical methods: modified ITT (participants who had baseline and at least 1 postrandomization primary efficacy measure, last observation carried forward)
Participants	N=290 participants with an ADHD diagnosis according to the DSM-IV-TR criteria Psychiatric comorbid conditions: NR Age range: 6-12 years Mean age (SD): 9 (1.8) Male: 201 (69%) ADHD subtype: Hyperactive-Impulsive: 12 (4%); Combined: 278 (96%)
Interventions	Four interventions: 1. Lisdexamphetamine (long-acting), 30 mg/day, N=71 2. Lisdexamphetamine (long-acting), 50 mg/day (30 mg/day for week 1 with forced dose escalation to 50 mg/day for weeks 2-4), N=74 3. Lisdexamphetamine (long-acting), 70 mg/day (30 mg/day for week 1 with forced-dose escalation to 50 mg/day for week 2 and 70 mg/day for weeks 3-4), N=73 4. Placebo, N=72 Duration: 28 days
Outcomes	Relevant outcomes: ADHD core symptom severity assessed with: ADHD-RS-IV, Conners' Parent Rating Scale-Revised: Short Form Clinical impression assessed with: CGI-Severity and CGI-Improvement Retention: proportion of participants who completed the trial Number of participants who experienced at least one adverse event Number of participants who dropped out due to any adverse event Adverse events
Other	Authors' affiliation: University Study funding: Pharmaceutical

Risk of bias assessment		
Bias	*Authors' judgement*	*Support for judgement*
Random sequence generation (selection bias)	Low	The sequence was generated by a computer program.
Allocation concealment (selection bias)	Unclear	Method of allocation concealment is not described.
Incomplete outcome data (attrition bias)	Low	Attrition was moderate (21%), however, 98% of randomized participants were included in the primary analysis. Reasons for drop-outs are provided.

Selective reporting (reporting bias)	Low	Data is provided on all outcomes listed in the protocol. Study appears to be free of selective reporting.
Other bias	High	Study was funded by industry and all authors are or have been affiliated with the pharmaceutical company that funded the study.
Blinding of participants and personnel (performance bias)	Low	Intervention and placebo are described as identical.
Blinding of outcome assessment (detection bias)	Unclear	Blinding of outcome assessment is not described.

5. Borcherding 1990	
Methods	Double-blind, single-center, placebo-controlled, crossover, randomized clinical trial Country: United States Statistical methods: ITT (all randomized participants were included in the analysis, with any missing data imputed with the group mean value)
Participants	N=46 participants with an ADHD diagnosis according to DSM-III criteria Psychiatric comorbid conditions: ODD, CD, reading developmental disorder, arithmetic disorder, dysthymic disorder Age range:6-12 years Mean age (SD): 8.6 (1.7) Male: 46 (100%) ADHD subtype: NR
Interventions	Three interventions: 1. Dextroamphetamine (short-acting), weight-based dosing increasing each week (children <30 kg received 10, 25, and 40 mg/day, bid; children >30 kg received 15, 30, and 45 mg/day. bid) 2. Methylphenidate hydrochloride, weight-based dosing increasing each week (children <30 kg received 25, 40, and 70 mg/day, bid; children >30 kg received 30, 50, and 90 mg/day, bid) 3. Placebo Duration: 63 days (three 21-day treatment periods)
Outcomes	Relevant outcomes: ADHD core symptom severity assessed with: Conners' Teacher Rating Scale and Conners' Parent Questionnaire Clinical impression assessed with: CGI-Improvement scale Academic performance assessed with: the Barnell Loft, Ltd Developing Key Concepts in Math test Retention: proportion of participants who completed the trial Adverse events Other outcomes:

	Nervous habits/mannerisms, compulsive acts and obsessive thinking assessed with Children's Psychiatric Rating scale Urine biochemistry Plasma biochemistry Renal clearance Continuous performance test	
Other	Authors' affiliation: National Institute of Mental Health Study funding: NR Outcomes were presented across 4 publications	
Risk of bias assessment		
Bias	*Authors' judgement*	*Support for judgement*
Random sequence generation (selection bias)	Unclear	Method of sequence generation is not described.
Allocation concealment (selection bias)	Unclear	Method of allocation concealment is not described.
Incomplete outcome data (attrition bias)	High	This study has 4 publications associated with it, and all reports have varying numbers of participants. Upon communication with the corresponding author of these reports, it was confirmed that the numbers of participants vary in the 4 publications due to missing data and drop-outs. Reasons for missing data were not provided.
Selective reporting (reporting bias)	High	The primary and secondary outcomes of this study are published in 4 different publications at different times. This means that although the methods of each of these reports are the same, the results sections are not consistent with what is stated in the methods.
Other bias	Unclear	No information on the validity of the primary outcome measure is provided.
Blinding of participants and personnel (performance bias)	Unclear	Blinding of participants and personnel is not described.
Blinding of outcome assessment (detection bias)	Unclear	Blinding of outcome assessment is not described.
6. Donnelly 1989		
Methods	Double-blind, single-center, placebo-controlled, crossover, randomized clinical trial Country: United States Statistical methods: ITT	
Participants	N=20 participants with an ADHD diagnosis according to DSM-III criteria	

	Psychiatric comorbid conditions: ODD, CD, mental learning disorder, language disorder Age range: NR Mean age (SD): 8 (2) Male: 20 (100%) ADHD subtype: NR
Interventions	Three interventions: 1. Dextroamphetamine (short-acting), weight based (0.5 mg/kg/day, bid) 2. Fenfluramine hydrochloride, weight-based dosing increasing each week (0.6mg/kg/day, 1.3 mg/kg/day, 2.0 mg/kg/day, bid) 3. Placebo Duration: 63 days (three 21-day treatment periods)
Outcomes	Relevant outcomes: ADHD core symptom severity assessed with: Conners' Abbreviated Teacher Rating Scale and Conners' Parent Questionnaire Clinical impression assessed with: CGI Adverse events Other outcomes: Children's Psychiatric Rating Scale Continuous performance test Motor activity Biochemical and platelet measures (urine and plasma) Measures of prolactin
Other	Authors' affiliation: university and National Institute of Mental Health Study funding: NR Donnelly 1986 is a pilot study of Donnelly 1989 and therefore has overlapping data. Unpublished data on the CGI were sought but not obtained

Risk of bias assessment

Bias	*Authors' judgement*	*Support for judgement*
Random sequence generation (selection bias)	Unclear	Method of sequence generation is not described.
Allocation concealment (selection bias)	Unclear	Method of allocation concealment is not described.
Incomplete outcome data (attrition bias)	Low	All participants recruited were included in the analyses. Only one drop-out, with reasons provided.
Selective reporting (reporting bias)	High	Data on majority of outcomes is missing (no means or measures of variance reported).
Other bias	Unclear	No information on the validity of the primary outcome measure is provided.

Blinding of participants and personnel (performance bias)	Unclear	Blinding of participants and personnel is not described.
Blinding of outcome assessment (detection bias)	Unclear	Blinding of outcome assessment is not described.
7. Findling 2011		
Methods	Double-blind, multi-center, placebo-controlled, parallel-group randomized clinical trial Country: United States Number of study sites: 45 Statistical methods: modified ITT (last observation carried forward)	
Participants	N=314 participants with an ADHD diagnosis according to DSM-IV-TR criteria Psychiatric comorbid conditions: NR Age range: 13-17 years Mean age (SD): 14.6 (1.31) Male: 249 (79%) ADHD subtype: Combined: 203 (65%)	
Interventions	Four interventions: 1. Lisdexamphetamine (long-acting), 30 mg/day, N=78* 2. Lisdexamphetamine (long-acting), 50 mg/day (30 mg/day for week 1 with forced dose escalation to 50 mg/day for weeks 2-4), N= 77* 3. Lisdexamphetamine (long-acting), 70 mg/day (30 mg/day for week 1 with forced-dose escalation to 50 mg/day for week 2 and 70 mg/day for weeks 3-4), N=78* 4. Placebo, N=79 Duration: 28 days	
Outcomes	Relevant outcomes: ADHD core symptom severity assessed with: ADHD-RS-IV Clinical impression assessed with: CGI-Severity and CGI-Improvement scales Quality of life assessed with: Youth Quality of Life-Research Version (YQOL-R) Retention: proportion of participants who completed the trial Number of participants who dropped out due to lack of efficacy Number of participants who experienced at least one adverse event Number of participants who dropped out due to any adverse event Adverse events	
Other	Authors' affiliation: university and pharmaceutical industry Study funding: pharmaceutical industry Registered at clinicaltrials.gov, ID: NCT00735371 *numbers are based on participants included in the safety analysis	
Risk of bias assessment		
Bias	*Authors' judgement*	*Support for judgement*

Random sequence generation (selection bias)	Low	Sequence was generated by a web-based computer system
Allocation concealment (selection bias)	Low	Allocation concealment was ensured using the web-based computer system and third party which serially numbered treatment bottles for each participant.
Incomplete outcome data (attrition bias)	Low	Attrition was moderate at 16%, however, 98% of randomized participants were included in the primary efficacy analysis. Reasons for drop-outs are provided.
Selective reporting (reporting bias)	Low	Data is provided on all outcomes listed in the protocol. Study appears to be free of selective reporting.
Other bias	High	Study was funded by industry and all authors are affiliated with the pharmaceutical company that funded the study.
Blinding of participants and personnel (performance bias)	Unclear	Blinding of participants and personnel is not described.
Blinding of outcome assessment (detection bias)	Unclear	Blinding of outcome assessment is not described.

8. Giblin 2011	
Methods	Double-blind, single-center, placebo-controlled, parallel-group randomized clinical trial Country: United States Statistical methods: modified ITT (all randomized participants who had both a baseline and a postrandomization primary outcome assessment)
Participants	N=24 participants with an ADHD diagnosis according to DSM-IV-TR criteria Psychiatric comorbid conditions: NR Age range: 6-12 years Mean age (SD): 9.65 (2.2) Male: 10 (42%) ADHD subtype: NR
Interventions	Four interventions: 1. Lisdexamphetamine (long-acting), 30 mg/day, N=3 2. Lisdexamphetamine (long-acting), 50 mg/day, N=11 3. Lisdexamphetamine (long-acting), 70 mg/day, N=2 4. Placebo, N=8 Duration: 28 days
Outcomes	Relevant outcomes: ADHD core symptom severity assessed with ADHD-RS-IV and Conners' Parent Rating Scale-Revised: Short form Clinical impression assessed with CGI-Severity scale

	Adverse events Other outcomes: Sleep onset latency assessed with polysomnography (PSG) Wake time after sleep onset assessed with PSG and actigraphy Number awakenings after sleep onset assessed with PSG Total sleep time assessed with PSG and actigraphy Sleep efficiency assessed with actigraphy	
Other	Authors' affiliation: Private organization and pharmaceutical industry Study funding: Pharmaceutical industry	
Risk of bias assessment		
Bias	*Authors' judgement*	*Support for judgement*
Random sequence generation (selection bias)	Unclear	Method of random sequence generation is not described.
Allocation concealment (selection bias)	Unclear	Method of allocation concealment is not described.
Incomplete outcome data (attrition bias)	Unclear	Incomplete outcome data is not addressed; number of completers is not reported
Selective reporting (reporting bias)	High	Study did not report means and any measures of variance for their primary outcome.
Other bias	High	Study was funded by industry. Authors are affiliated with industry.
Blinding of participants and personnel (performance bias)	Unclear	Blinding of participants and personnel is not described.
Blinding of outcome assessment (detection bias)	Unclear	Blinding of outcome assessment is not described.
9. Gillberg 1997		
Methods	Double-blind, multi-center, placebo-controlled, parallel group, randomized clinical trial Country: Sweden Number of sites: 4 Statistical methods: Modified ITT (for inclusion into the analysis at least 2 measurements had to be available, with the last observation carried forward)	
Participants	N=62 participants with an ADHD diagnosis according to DSM-III-R criteria Psychiatric comorbid conditions: ODD, CD, anxiety, autistic disorder, pervasive development disorder, motor tic disorder, Tourette syndrome, mild mental retardation Age range: 6-11 years Mean age (SD): 9 (1.6) Male: 52 (84%)	

	ADHD subtype: NR	
Interventions	Two interventions: 1. Mixed amphetamine salts (short-acting), dosage was titrated from 10 mg/day bid to a maximum of 60 mg/day bid, N=32 2. Placebo, N=30 Duration: 365 days	
Outcomes	Relevant outcomes: ADHD core symptom severity assessed with: Conners' Teacher Rating Scale and Conners' Parent Rating Scale Academic performance assessed with: Wechsler Intelligence Scale for Children Adverse events Other outcomes: Depression assessed with the Birleson Depression Self-report Scale Mood assessed with: the McGrath Test	
Other	Authors' affiliation: University and pharmaceutical industry Study funding: Pharmaceutical industry and public funds	
Risk of bias assessment		
Bias	*Authors' judgement*	*Support for judgement*
Random sequence generation (selection bias)	Unclear	Method of random sequence generation is not described.
Allocation concealment (selection bias)	Unclear	Method of allocation concealment is not described.
Incomplete outcome data (attrition bias)	Low	Attrition was moderate at 14%, however, all drop-outs were prior to randomization. Reasons for drop-out are provided. All individuals randomized were included in the primary analysis.
Selective reporting (reporting bias)	High	Data on two outcomes (Birleson Depression Self-report scale and McGrath Test) listed in the methods section is not reported.
Other bias	High	Study was funded by industry and authors are affiliated with the pharmaceutical company that funded the study.
Blinding of participants and personnel (performance bias)	Unclear	Blinding of participants and personnel is not described.
Blinding of outcome assessment (detection bias)	Unclear	Blinding of outcome assessment is not described.
10. James 2001		
Methods	Double-blind, placebo-controlled, crossover, randomized clinical trial	

	Country: United States Statistical methods: per protocol
Participants	N=35 participants with an ADHD diagnosis according to DSM-IV criteria Psychiatric comorbid conditions: ODD, anxiety, enuresis, dysthymic disorder, learning disorder Age range: 6.9-12.2 years Mean age (SD): 9.1 (1.5) Male: 21 (60%) ADHD subtype: Combined: 35 (100%)
Interventions	Four interventions: 1. Dextroamphetamine (short-acting)* 2. Dextroamphetamine (long-acting)*,** 3. Placebo *doses were individualized and based on age, weight, prior medication experience and symptom severity (overall mean dose range: 7.8-12.8 mg/day) Duration: 56 days (four 14-day treatment periods)
Outcomes	Relevant outcomes: ADHD core symptom severity assessed with: Conners' Teacher Rating Scale, Conners' Parent Behavior Rating Scale Academic performance assessed with: 5-minute timed math task Adverse events Other outcomes: Motor activity assessed with an actometer
Other	Authors' affiliation: National Institute of Mental Health Study funding: NR **Dextroamphetamine long-acting was randomly chosen to represent the amphetamine group in this study for binary outcomes

Risk of bias assessment

Bias	*Authors' judgement*	*Support for judgement*
Random sequence generation (selection bias)	Unclear	Method of random sequence generation is not described.
Allocation concealment (selection bias)	Unclear	Method of allocation concealment is not described.
Incomplete outcome data (attrition bias)	Low	Attrition was low at 7%, and reasons were provided. All drop-outs occurred prior to randomization. All participants who were randomized completed the trial and were included in the primary analysis.
Selective reporting (reporting bias)	Low	Data is provided on all outcomes listed in the methods. Study appears to be free of selective reporting.

Other bias	Low	Study appears to be free of other biases
Blinding of participants and personnel (performance bias)	Low	Intervention and placebo are described as identical
Blinding of outcome assessment (detection bias)	Unclear	Blinding of outcome assessment is not described.
11. Manos 1999		
Methods	Double-blind, single-center, placebo-controlled, crossover randomized clinical trial Country: United States Statistical methods: unclear	
Participants	N=84 participants with an ADHD diagnosis according to DSM-IV criteria Psychiatric comorbid conditions: ODD, anxiety, mood disorder, learning disability Age range: 5-17 years Mean age (SD): 10.1 (NR) Male: 66 (79%) ADHD subtypes: Inattententive: 38 (45%), Combined: 46 (55%)	
Interventions	1. Mixed amphetamine salts (short-acting) group, N=42: i. 5 mg/day, qd ii. 10 mg/day, qd iii. 15 mg/day, qd iv. Placebo 2. Methylphenidate group, N=42: i. 5 mg/day, bid ii. 10 mg/day, bid iii. 15 mg/day, bid iv. Placebo Subjects received either the 4 methylphenidate OR adderall conditions (determined by participant's physician) in a randomly assigned sequence Duration: 28 days (four 7-day treatment periods)	
Outcomes	Relevant outcomes: ADHD core symptom severity assessed with: ADHD Rating Scale. Abbreviated Symptoms Questionnaire-Parent, Abbreviated Symptoms Questionnaire-Teacher Adverse events Other outcomes: Concentration in school assessed with: School Situations Questionnaire-Revised	
Other	Authors' affiliations: university Study funding: Public funds	

Risk of bias assessment		
Bias	*Authors' judgement*	*Support for judgement*
Random sequence generation (selection bias)	Unclear	Method of random sequence generation is not described.
Allocation concealment (selection bias)	Low	A third party pharmacist prepared individually sealed bottles dated by week.
Incomplete outcome data (attrition bias)	Unclear	Study did not describe if any participants dropped out from the trial.
Selective reporting (reporting bias)	Low	Data is provided on all outcomes listed in the methods. Study appears to be free of selective reporting.
Other bias	Low	Study appears to be free of other biases
Blinding of participants and personnel (performance bias)	Low	Intervention and placebo are described as identical.
Blinding of outcome assessment (detection bias)	Low	Clinician, teacher, and parent (outcome assessors) were blind to treatment order.

12. McCracken 2003	
Methods	Double-blind, multi-center, placebo-controlled, crossover, randomized clinical trial Country: United States Number of study sites: 4 Statistical methods: modified ITT
Participants	N=51 participants with an ADHD diagnosis according to the DSM-IV criteria Psychiatric comorbid conditions: NR Age range: 6-12 years Mean age (SD): 9.5 (1.9) Male: 44 (86%) ADHD subtypes: Hyperactive-Impulsive: 1 (2%); Combined: 50 (98%)
Interventions	Five interventions: 1. Mixed amphetamine salts (short-acting), 10 mg qd 2. Mixed amphetamine salts (long-acting), 10 mg qd 3. Mixed amphetamine salts (long-acting), 20 mg qd 4. Mixed amphetamine salts (long-acting), 30 mg qd 5. Placebo Duration: 35 days (five 7-day treatment periods)
Outcomes	Relevant outcomes: ADHD core symptom severity assessed with: SKAMP-attention subscale Academic performance assessed with: PERMP test Retention: Proportion of participants who completed the trial Number of participants who dropped out due to any adverse event

	Adverse events Other outcomes: Conduct problems assessed with: SKAMP-deportment subscale	
Other	Authors' affiliation: University Study funding: Pharmaceutical industry *Mixed amphetamine salts, 20 mg/day was randomly chosen to represent the amphetamine group in this study for binary outcomes	
Risk of bias assessment		
Bias	*Authors' judgement*	*Support for judgement*
Random sequence generation (selection bias)	Unclear	Method of random sequence generation is not described.
Allocation concealment (selection bias)	Unclear	Method of allocation concealment is not described
Incomplete outcome data (attrition bias)	Low	Attrition was low (4%) and reasons were provided. All randomized participants were included in the primary analysis.
Selective reporting (reporting bias)	High	Methods state that parent, teacher and clinician/research staff would complete a side effect rating scale, however, data are only presented on that completed by the parent.
Other bias	High	Study was funded by industry and most authors are affiliated with the pharmaceutical company that funded the study.
Blinding of participants and personnel (performance bias)	Unclear	Blinding of participants and personnel is not described
Blinding of outcome assessment (detection bias)	Unclear	Blinding of outcome assessment is not described
13. Nemzer 1986		
Methods	Double-blind, placebo-controlled, crossover, randomized clinical trial Country: United States Statistical methods: NR	
Participants	N=14 participants with an ADHD diagnosis according to DSM-III criteria Psychiatric comorbid conditions: NR Age range: 7-12 years Mean age (SD): 9.36 (NR) Male: 11 (79%) ADHD subtype: NR	
Interventions	Four interventions: 1. Dexedrine (short-acting), weight-based dosing (children <32 kg	

	received 5 mg/day, bid; children >32 kg received 10 mg/day, bid) 2. Tyrosine supplement, 140 mg/kg/day 3. Tryptophan supplement, 100 mg/kg/day 4. Placebo Duration: 28 days (four 7-day treatment periods)		
Outcomes	Relevant outcomes: ADHD core symptom severity assessed with: Conners' Parent Questionnaire and Conners' Teacher Rating Scale Academic performance assessed with: WISC-R Adverse events Other outcomes: Quay-Peterson Behavior Checklist David's Hyperkinetic Scale Tyrosine serum levels Tryptophan serum levels		
Other	Authors' affiliation: university Study funding: NR		
Risk of bias assessment			
Bias	*Authors' judgement*	*Support for judgement*	
Random sequence generation (selection bias)	Unclear	Method of random sequence generation is not described.	
Allocation concealment (selection bias)	Unclear	Method of allocation concealment is not described	
Incomplete outcome data (attrition bias)	Unclear	Study does not address drop-outs.	
Selective reporting (reporting bias)	Low	Data is provided on all outcomes listed in the methods. Study appears to be free of selective reporting.	
Other bias	Low	Study appears to free of other biases	
Blinding of participants and personnel (performance bias)	Low	Intervention and placebo are described as identical.	
Blinding of outcome assessment (detection bias)	Unclear	Blinding of outcome assessment is not described	
14. Pliszka 2000			
Methods	Double-blind, placebo-controlled, parallel group, randomized clinical trial Country: United States Statistical methods: modified ITT (last observation carried forward)		
Participants	N=58* participants with an ADHD diagnosis according to the Diagnostic Interview Schedule for Children Psychiatric comorbid disorders: ODD, CD, anxiety Age range:6-10 years Mean age (SD): 8.2 (1.6)		

	Male: NR ADHD subtype: NR
Interventions	Three interventions: 1. Mixed amphetamine salts (short-acting), weight-based, titrated dosing (maximum dose for children <60 lbs was 15 mg/day; maximum daily dose for children >60 lbs was 30 mg/day), N=20 2. Methylphenidate, weight-based, titrated dosing (maximum dose for children <60 lbs was 25 mg/day; maximum daily dose for children >60 lbs was 50 mg/day), N=20 3. Placebo, N=18 Duration: 21 days
Outcomes	Relevant outcomes: ADHD core symptom severity assessed with: IOWA Conners' Teacher Rating Scale and Conners' Parent Global Index Clinical impression: CGI-Improvement Number of participants who dropped out due to any adverse event Adverse events Other outcomes: Aggression/defiance assessed with Conners' Teacher Rating Scale
Other	Author's affiliation: University and pharmaceutical industry Study funding: pharmaceutical industry *One participant dropped out before the end of the study, and was not accounted for in the participant characteristic description (data only provided on 58 participants)

Risk of bias assessment		
Bias	*Authors' judgement*	*Support for judgement*
Random sequence generation (selection bias)	Unclear	Method of random sequence generation is not described.
Allocation concealment (selection bias)	Unclear	Method of allocation concealment is not described
Incomplete outcome data (attrition bias)	Low	Attrition was low (10%) and reasons are provided. All participants who were randomized were included in the primary analysis.
Selective reporting (reporting bias)	Low	Data is provided on all outcomes listed in the methods. Study appears to be free of reporting bias.
Other bias	High	Study was funded by industry and all authors are affiliated with the pharmaceutical company that funded the study.
Blinding of participants and personnel (performance bias)	Low	Intervention and placebo are described as identical.

Blinding of outcome assessment (detection bias)	Unclear	Blinding of outcome assessment is not described
15. Sharp 1999		
Methods	Double-blind, placebo-controlled, crossover, randomized clinical trial Country: United States Statistical methods: ITT (missing data were imputed using group means)	
Participants	N=32* participants with an ADHD diagnosis according to DSM-III-R/DSM-IV criteria Psychiatric comorbid disorders: ODD, CD, depression, separation anxiety, specific phobias, tic disorder, enuresis, reading disorder Age range:6.2-12.7 years Mean age (SD): 8.9 (1.7) Male: 0 (0%) ADHD subtype: combined: 32 (100%)	
Interventions	Three interventions: 1. Dextroamphetamine (short-acting), weight-based dosing, and increasing over time (mean doses of 0.23 mg/kg/day, 0.43 mg/kg/day, and 0.64 mg/kg/day, bid for weeks 1, 2, and 3 respectively) 2. Methylphenidate, weight-based dosing, and increasing over time (mean doses of 0.45 mg/kg/day, 0.85 mg/kg/day, and 1.28 mg/kg/day, qd for weeks 1, 2, and 3 respectively) 3. Placebo Duration: 63 days (three 21-day treatment periods)	
Outcomes	Relevant outcomes: ADHD core symptom severity assessed with: Parent Conners' Rating Scale and Teacher Conners' Rating Scale Clinician impression assessed with: CGI-Severity and CGI-Improvement scales Retention: proportion of participants who completed the trial Adverse events	
Other	Authors' affiliation: university and National Institute for Mental Health Study funding: NR Unpublished data on the ADHD core symptoms were sought but were not obtained. *Clinical characteristics were presented on 42 participants, 10 of which participated in a separate pilot program not related to the study.	
Risk of bias assessment		
Bias	*Authors' judgement*	*Support for judgement*
Random sequence generation (selection bias)	Unclear	Method of random sequence generation is not described.

Allocation concealment (selection bias)	Unclear	Method of allocation concealment is not described
Incomplete outcome data (attrition bias)	Low	Attrition was moderate (14%) and reasons are provided. All participants who were randomized were included in the primary analysis.
Selective reporting (reporting bias)	Low	Data is provided on all outcomes listed in the methods. Study appears to be free of reporting bias.
Other bias	Low	Study appears to be free of other biases.
Blinding of participants and personnel (performance bias)	Low	Intervention and placebo are described as identical.
Blinding of outcome assessment (detection bias)	Unclear	Blinding of outcome assessment is not described

16. Shekim 1986	
Methods	Double-blind, placebo-controlled, crossover, randomized clinical trial Country: United States Statistical methods: NR
Participants	N=22 participants with an ADHD diagnosis according to DSM-III criteria Psychiatric comorbid disorders: None Age range: 6-12 Mean age (SD): 9.75 (2.08) Male: 22 (100%) ADHD subtype: NR
Interventions	Two interventions: 1. Dextroamphetamine (short-acting), weight-based at 0.3 mg/kg, bid, and titrated upwards during trial period 2. Placebo Duration: 28 days (two 14-day treatment periods)
Outcomes	Relevant outcomes Academic performance assessed with: Wide Range Achievement Test-Math subset Other outcomes Wide Range Achievement Test- spelling and reading subsets Monoamine oxidase activity Continuous performance test reaction time
Other	Authors' affiliation: university Study funding: public funds.

Risk of bias assessment		
Bias	*Authors' judgement*	*Support for judgement*

Random sequence generation (selection bias)	Unclear	Method of random sequence generation is not described.
Allocation concealment (selection bias)	Unclear	Method of allocation concealment is not described
Incomplete outcome data (attrition bias)	High	Reasons for exclusions from the analysis are not provided. Methods of analysis are not described. Number of individuals included in the analyses is not reported.
Selective reporting (reporting bias)	Low	Data is provided on all outcomes listed in the methods. Study appears to be free of reporting bias.
Other bias	Low	Study appears to be free of other biases.
Blinding of participants and personnel (performance bias)	Unclear	Blinding of participants and personnel is not described.
Blinding of outcome assessment (detection bias)	Unclear	Blinding of outcome assessment is not described

17. Short 2004	
Methods	Double-blind, placebo-controlled, crossover, randomized clinical trial Country: United States Statistical methods: per protocol
Participants	N=34* participants with an ADHD diagnosis according to DSM-IV criteria Psychiatric comorbid conditions: NR Age range: 3-5.9 years old Mean age (SD): 5.3 (NR) Male: 24 (85%) ADHD subtype: Inattentive: 5 (17%); Hyperactive-impulsive and combined: 23 (83%)
Interventions	2 groups: Amphetamine group: 1. Mixed amphetamine salts (short-acting), 5 mg/day qd 2. Mixed amphetamine salts (short-acting), 10 mg/day qd 3. Mixed amphetamine salts (short-acting), 15 mg/day qd 4. Placebo Methylphenidate group: 1. Methylphenidate, 5 mg/day bid 2. Methylphenidate, 10 mg/day bid 3. Methylphenidate, 15 mg/day bid 4. Placebo Amphetamine or methylphenidate was determined by a physician. Duration: 28 days (four 7-day treatment periods)
Outcomes	Relevant outcomes:

	ADHD core symptom severity assessed with: Conners Abbreviated Symptoms Questionnaire (teacher and parent); ADHD-RS-IV; Home Situations Questionnaire Adverse events	
Other	Authors' affiliations: University Study funding: Public funds The authors did not separate the two active interventions (amphetamine and methylphenidate) in their analysis. Authors were contacted on 3 occasions to obtain the data on amphetamines only, but we received no response. *Clinical characteristics are only presented on participants who were included in the analysis (N=28)	
Risk of bias assessment		
Bias	*Authors' judgement*	*Support for judgement*
Random sequence generation (selection bias)	Unclear	Method of random sequence generation is not described.
Allocation concealment (selection bias)	Unclear	Method of allocation concealment is not described
Incomplete outcome data (attrition bias)	High	Report does not describe reasons for attrition. In addition 6 participants were dropped from the analysis, and their last data point was not carried forward.
Selective reporting (reporting bias)	High	Authors assessed 2 active interventions (amphetamines and methylphenidate) compared with placebo, however, they did not separate the two interventions out in the analysis, therefore one could not decipher the difference in efficacy between the two stimulant medications.
Other bias	Low	Study appears to be free of other biases.
Blinding of participants and personnel (performance bias)	Low	Intervention and placebo are described as identical.
Blinding of outcome assessment (detection bias)	Low	Outcome assessors were blind to order of trial interventions.
18. Spencer 2006		
Methods	Double-blind, multi-center, placebo-controlled, parallel-group, randomized clinical trial Country: United States Number of study sites: NR Statistical methods: modified ITT (those with at least one post-baseline primary efficacy assessment)	
Participants	N=287* participants with an ADHD diagnosis according to the DSM-IV-TR criteria Psychiatric comorbid conditions: NR	

	Age range: 13-16 years Mean age (SD): 14.2 (1.2) Male: 182 (66%) ADHD subtypes: Inattentive: 114 (41%); hyperactive-impulsive: 7 (2.5%); combined: 157 (56.5%)
Interventions	Five interventions: 1. Mixed amphetamine salts (long-acting), 10 mg/day qd, N=56 2. Mixed amphetamine salts (long-acting), 20 mg/day qd (10 mg/day for week 1 with forced-dose escalation to 20 mg/day for weeks 2-4), N=56 3. Mixed amphetamine salts (long-acting), 30 mg/day qd (10 mg/day for week 1, with forced-dose escalation to 20 mg/day for week 2, and 30 mg/day for weeks 3-4), N=58 4. Mixed amphetamine salts (long-acting), 40 mg/day qd (10 mg/day for week 1, with forced-dose escalation to 20 mg/day for week 2, 30 mg/day for week 3, and 40 mg/day for week 4), N= 63 5. Placebo, N=54 Duration: 28 days (four 7-day treatment periods)
Outcomes	Relevant outcomes ADHD core symptom severity assessed with ADHD-RS-IV Clinical impression assessed with: CGI-Severity and CGI-Improvement scales Retention: proportion of participants who completed the trial Number of participants who dropped out due to any adverse event Adverse events Other outcomes: Vital signs Body weight
Other	Clinicaltrials.gov identifier: NCT00507065 Authors' affiliation: University and pharmaceutical industry Study funding: Pharmaceutical industry *clinical characteristics are only provided on the study's intent-to-treat population (N=278)

Risk of bias assessment		
Bias	*Authors' judgement*	*Support for judgement*
Random sequence generation (selection bias)	Unclear	Method of random sequence generation is not described.
Allocation concealment (selection bias)	Unclear	Method of allocation concealment is not described.
Incomplete outcome data (attrition bias)	Low	Attrition was low at 10%, and reasons were provided. Only 3% of participants were excluded from the analysis due to no post-baseline primary efficacy assessment.

Selective reporting (reporting bias)	Low	Data is provided on all outcomes listed in the protocol. Study appears to be free of selective reporting.
Other bias	High	Study was funded by industry and most authors are affiliated with the pharmaceutical company that funded the study.
Blinding of participants and personnel (performance bias)	Unclear	Blinding of participants and personnel is not described.
Blinding of outcome assessment (detection bias)	Unclear	Blinding of outcome assessment is not described.

19. Swanson 1998	
Methods	Double-blind, placebo-controlled, crossover, randomized clinical trial Country: United States Statistical methods: unclear
Participants	N=33 participants with an ADHD diagnosis according to DSM-IV criteria Psychiatric comorbid conditions: NR Age range: 7-14 years Mean age (SD): 10.58 (1.81) Male: 26 (79%) ADHD subtypes: NR
Interventions	Six interventions: 1. Mixed amphetamine salts (short-acting), 5 mg/day qd 2. Mixed amphetamine salts (short-acting), 10 mg/day qd 3. Mixed amphetamine salts (short-acting), 15 mg/day qd 4. Mixed amphetamine salts (short-acting), 20 mg/day qd 5. Methylphenidate (dose determined by physician) 6. Placebo Duration: 49 days (six 7-day treatment periods defined by the six medication conditions above, plus an extra 7-day period to provide an opportunity to make up any missed weeks)
Outcomes	Relevant outcomes: ADHD core symptom severity assessed with: SKAMP-attention subscale Academic performance assessed with: PERMP Retention: proportion of randomized participants who completed the trial Other outcomes: Conduct problems assessed with: SKAMP-deportment subscale
Other	Authors' affiliation: university and pharmaceutical industry Study funding: Pharmaceutical industry

	Unpublished data on outcomes were sought but not obtained (ADHD core symptom severity and academic performance)	
Risk of bias assessment		
Bias	*Authors' judgement*	*Support for judgement*
Random sequence generation (selection bias)	Unclear	Method of random sequence generation is not described.
Allocation concealment (selection bias)	Unclear	Method of allocation concealment is not described.
Incomplete outcome data (attrition bias)	High	Although attrition was low (8%), reasons for drop-out were not provided, and only 88% of participants contributed to the primary analysis.
Selective reporting (reporting bias)	Low	Data is provided on all outcomes listed in the methods. Study appears to be free of selective reporting.
Other bias	High	Study was funded by industry.
Blinding of participants and personnel (performance bias)	Low	Intervention and placebo are described as identical.
Blinding of outcome assessment (detection bias)	Unclear	Blinding of outcome assessment is not described.
20. Wigal 2009		
Methods	Double-blind, multi-center, placebo-controlled, crossover, randomized clinical trial Country: United States Number of study sites: 7 Statistical methods: Modified ITT (participants who received at least one dose of study medication, with at least 1 post-randomization measurement of the primary efficacy variable-last observation carried forward)	
Participants	N=117* participants with an ADHD diagnosis according to DSM-IV-TR criteria Psychiatric comorbid disorders: NR Age range: 6-12 Mean age (SD): 10.1 (1.5) Male: 98 (76%) ADHD subtypes: NR	
Interventions	Two interventions: 1. Lisdexamphetamine (long-acting), participants were titrated to their optimal dose (30 mg/day (n=58); 50 mg/day (n=50); 70 mg/day (n=21)) 2. Placebo Duration: 14 days (two 7-day treatment periods)	
Outcomes	Relevant outcomes:	

	ADHD core symptom severity assessed with: SKAMP scale, SKAMP attention subscale, ADHD-RS-IV Academic performance assessed with: PERMP Clinical impression assessed with: CGI-Severity and CGI-Improvement scales Retention: number of participants who completed the study Number of participants who experienced at least one adverse event Number of participants who dropped out due to an adverse event Adverse events
Other	Authors' affiliation: university and pharmaceutical industry Study funding: pharmaceutical industry clinicaltrials.gov ID: NCT00500149 *Clinical characteristics are provided on 129 participants who were first enrolled into an open-label dose-optimisation phase, of these, 117 participants were randomized to the double-blind phase

Risk of bias assessment		
Bias	*Authors' judgement*	*Support for judgement*
Random sequence generation (selection bias)	Unclear	Method of random sequence generation is not described.
Allocation concealment (selection bias)	Unclear	Method of allocation concealment is not described.
Incomplete outcome data (attrition bias)	Low	Attrition was low (9%) and reasons were provided. 97% of randomized participants were included in the primary analysis. Four individuals were not included due to no post-baseline efficacy measure.
Selective reporting (reporting bias)	Unclear	Additional outcomes not included in the registered protocol are reported on.
Other bias	High	Study was funded by industry and all authors are affiliated with the pharmaceutical company that funded the study.
Blinding of participants and personnel (performance bias)	Low	Intervention and placebo are described as identical.
Blinding of outcome assessment (detection bias)	Unclear	Blinding of outcome assessment is not described.

Table 3.11.2 Characteristics of excluded studies

Study	Reason for Exclusion
Akhondzadeh 2003	No placebo comparison assessed
Alexandris 1968	ADHD diagnosis not confirmed using formal diagnostic criteria
Arnold 2011	No placebo-amphetamine comparison
Biederman 2006	Not a randomized controlled trial
Biederman 2008	Not a randomized controlled trial
Boellner 2010	Not a randomized controlled trial
Brown 2010	Not a randomized controlled trial
Denhoff 1971	ADHD diagnosis not confirmed using formal diagnostic criteria
Donner 2008	Not a randomized controlled trial
Efron 1997	No placebo control used
Findling 2009	Not a randomized controlled trial
Greenhill 2003	No placebo control used
Kamien 1998	Study design was a series of multiple crossover randomized controlled trials (N-of-1 trials); our review included on single crossover randomized controlled trials
Lopez 2008	Not a randomized controlled trial
McGough 2005	Not a randomized controlled trial
Najib 2009	Not a randomized controlled trial
Nikles 2006	Study design was a series of multiple crossover randomized controlled trials (N-of-1 trials); our review included on single crossover randomized controlled trials
Quintana 2007	Not a randomized controlled trial
Sleator 1974	Not a randomized controlled trial
Spencer 2005	Not a randomized controlled trial
Spencer 2006b	ADHD diagnosis not confirmed using formal diagnostic criteria
Turgay 2010	Not a randomized controlled trial
Wigal 2009b	Not a randomized controlled trial
Wigal 2010a	Not a randomized controlled trial
Wigal 2010b	Study participants were adults

Table 3.11.3 Summary of Findings Table

Amphetamines compared to placebo for attention deficit/hyperactivity disorder in children and adolescents						
Patient or population: Children or adolescents with ADHD **Settings:** **Intervention:** Amphetamines (i.e. dexamphetamine, lisdexamphetamine, mixed amphetamine salts) **Control:** Placebo						
Outcomes	**Illustrative comparative risks* (95% CI)**		**Relative effect (95%CI)**	**No. of participants (studies)**	**Quality of the evidence (GRADE)**	**Comments**
	Assumed risk	**Corresponding risk**				
	Placebo	**Amphetamine**				
Total ADHD symptom score-Parent ratings		The mean total ADHD symptom score-parent ratings in the intervention group was **0.4 standard deviations lower** (-0.53 to -0.27)		1046 (6 studies)	⊕⊕⊕⊖ **moderate**[1]	SMD -0.4 (-0.53, -0.27)
Total ADHD symptom score-Teacher ratings		The mean total ADHD symptom score-teacher ratings in the intervention group was **0.55 standard deviations lower** (-0.83 to -0.27)		745 (5 studies)	⊕⊕⊕⊕ **high**	SMD -0.55 (-0.83, -0.27)
Total ADHD symptom score-Clinician ratings		The mean total ADHD symptom score-clinician ratings in the intervention group was **0.84 standard deviations lower** (-1.32 to -0.36)		813 (3 studies)	⊕⊕⊖⊖ **low**[1,2]	SMD -0.84 (-1.32, -0.36)
Proportion of responders	**190 per 1000**	**593 per 1000**	**RR 3.12** (2.32 to 4.20)	1997 (8 studies)	⊕⊖⊖⊖ **very low**[1,2,3,4]	
Academic performance		The mean academic performance score in the intervention group was **0.51 standard deviations higher** (0.31 to 0.70)		632 (7 studies)	⊕⊕⊕⊖ **moderate**[4]	SMD 0.51 (0.31, 0.70)

Retention: proportion of participants who completed the trial	**807 per 1000**	**844 per 1000**	**RR 1.05** (0.99 to 1.11)	1946 (8 studies)	⊕⊖⊖⊖ **very low**[1,2,4]	
Proportion of participants who experienced at least 1 adverse event	**420 per 1000**	**628 per 1000**	RR 1.31 (1.16 to 1.48)	1548 (5 studies)	⊕⊕⊖⊖ **low**[1,4]	

*The basis for the **assumed risk** (e.g. the median control group risk across studies) is provided in the footnotes. The **corresponding risk** (and its 95% confidence interval) is based on the assumed risk in the comparison group and the **relative effect** of the intervention (and its 95% CI).
ADHD: Attention deficit/hyperactivity disorder; **CI:** Confidence interval; **RR:** risk ratio

GRADE Working Group grades of evidence
High quality: further research is very unlikely to change our confidence in the estimate of effect
Moderate quality: Further research is likely to have an important impact on our confidence interval in the estimate of effect and may change the estimate
Low quality: Further research is very likely to have an important impact on our confidence estimate of effect and is likely to change the estimate
Very low quality: We are very uncertain about the estimate

[1]The majority of studies included in this comparison were industry funded

[2]High statistical heterogeneity was found

[3]Wide 95% confidence interval indicates that the intervention effect for this outcome is highly variable

[4]This comparison includes three different types of amphetamine derivatives

3.12 FIGURES

Figure 3.12.1 Study flow

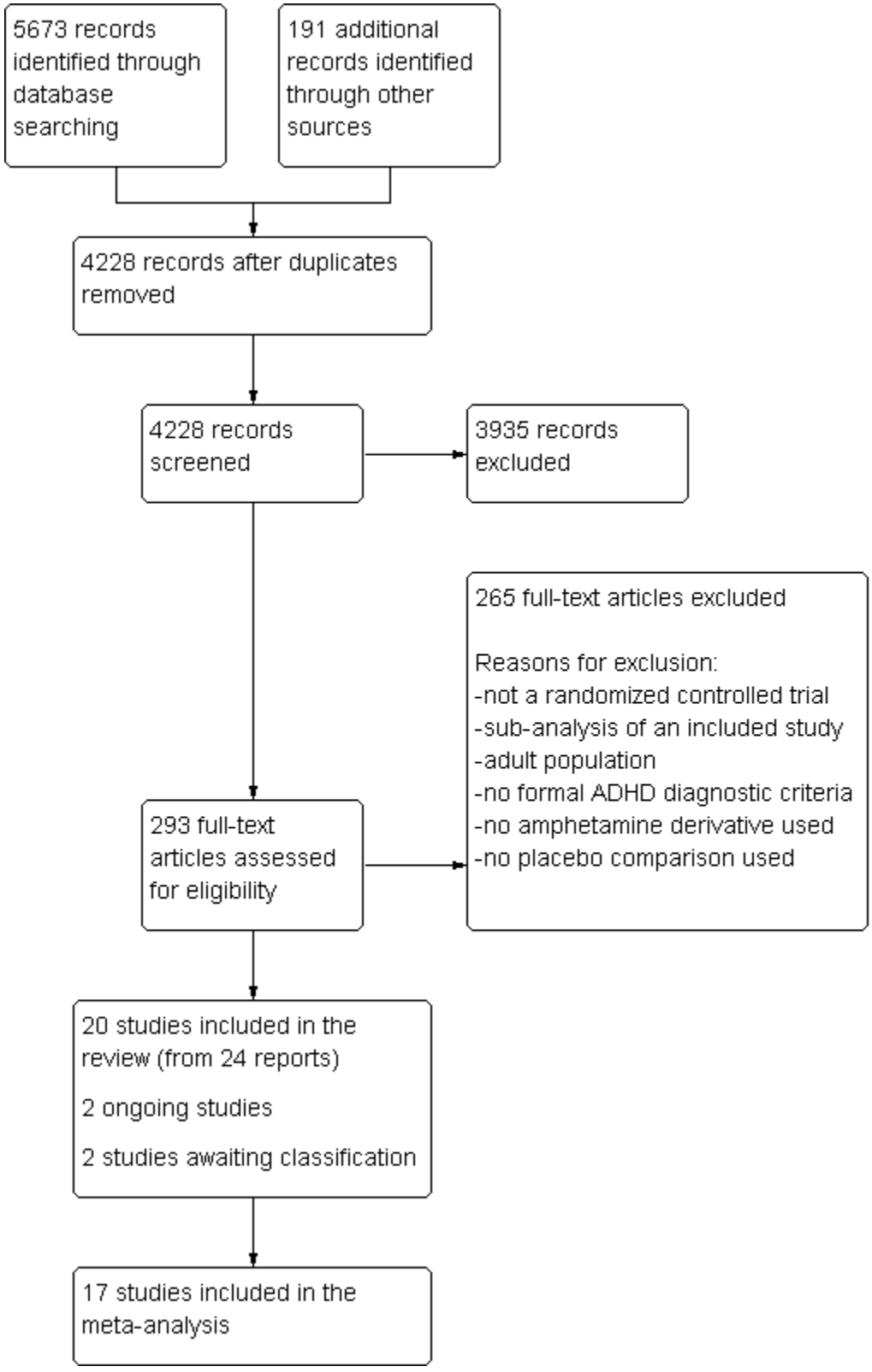
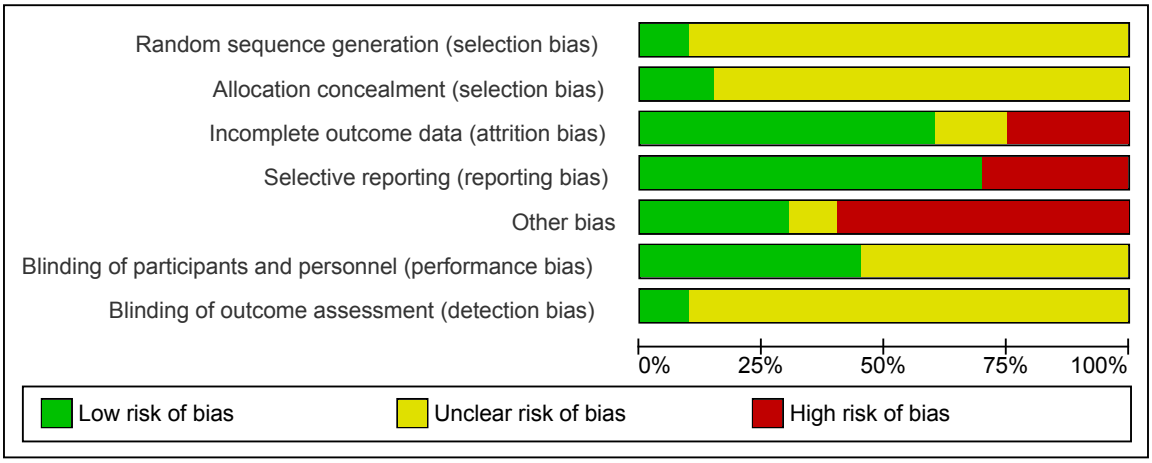


Figure 3.12.2 Risk of bias graph



Review authors' judgements about each risk of bias item presented as percentages across all included studies.

3.12.3 Primary Analysis Figures

Figure 3.12.3.1: Analysis 1.1 Total ADHD symptom score-Parent ratings

Study or Subgroup	Amphetamine Mean	SD	Total	Placebo Mean	SD	Total	Weight	Std. Mean Difference IV, Random, 95% CI
Barkley 2000	20.15	8.95	31	21.9	12.5	31	6.8%	-0.16 [-0.66, 0.34]
Biederman 2002	7.8	10.7	360	11.8	8.8	203	55.9%	-0.40 [-0.57, -0.22]
Biederman 2007b	18.6	59.8	213	34.3	34.8	72	23.4%	-0.29 [-0.55, -0.02]
Manos 1999	11.79	9.86	42	20.01	11.68	42	8.6%	-0.75 [-1.20, -0.31]
Nemzer 1986	13.29	6.4	14	17.21	6.2	14	2.9%	-0.60 [-1.36, 0.16]
Pliszka 2000	1.04	0.65	12	1.54	0.88	12	2.5%	-0.62 [-1.45, 0.20]
Total (95% CI)			**672**			**374**	**100.0%**	**-0.40 [-0.53, -0.27]**

Heterogeneity: Tau² = 0.00; Chi² = 4.59, df = 5 (P = 0.47); I² = 0%
Test for overall effect: Z = 6.00 (P < 0.00001)

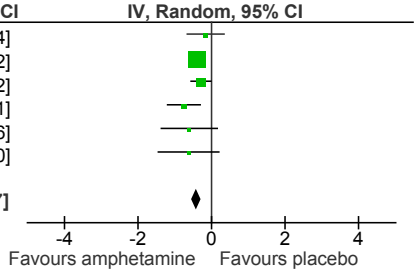


Figure 3.12.3.2: Analysis 1.2 Parent ratings of hyperactivity/impulsivity

Study or Subgroup	Amphetamine Mean	SD	Total	Placebo Mean	SD	Total	Weight	Std. Mean Difference IV, Random, 95% CI
Biederman 2007b	-12.2	12.84	213	-3.4	12.84	72	61.9%	-0.68 [-0.96, -0.41]
Borcherding 1990	0.8	1.87	31	1.75	1.87	31	18.0%	-0.50 [-1.01, 0.00]
James 2001	59.6	14.5	35	68	14.5	35	20.1%	-0.57 [-1.05, -0.09]
Total (95% CI)			**279**			**138**	**100.0%**	**-0.63 [-0.84, -0.41]**

Heterogeneity: Tau² = 0.00; Chi² = 0.45, df = 2 (P = 0.80); I² = 0%
Test for overall effect: Z = 5.74 (P < 0.00001)

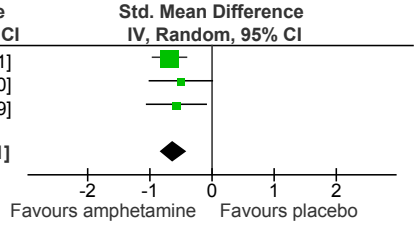


Figure 3.12.3.3: Analysis 1.3 Parent ratings of inattention

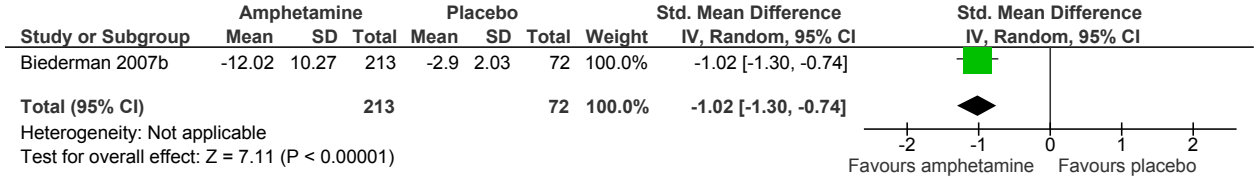


Figure 3.12.3.4: Analysis 1.4 Total ADHD symptom score-Teacher ratings

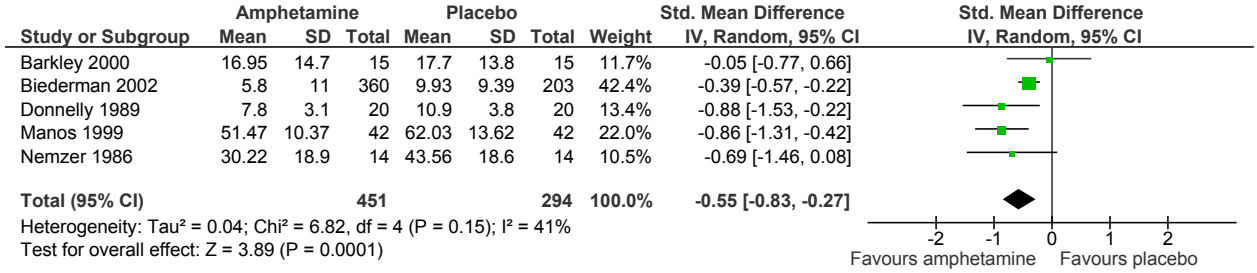


Figure 3.12.3.5: Analysis 1.5 Teacher ratings of hyperactivity/impulsivity

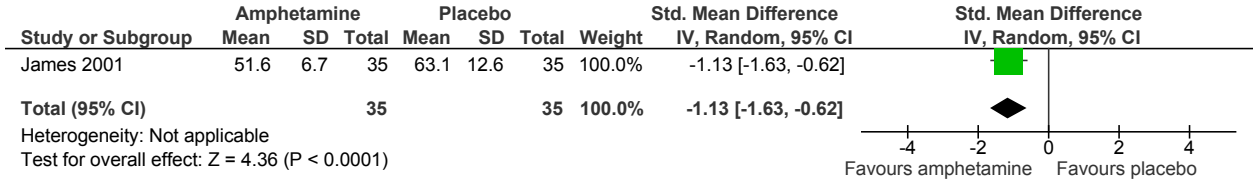


Figure 3.12.3.6: Analysis 1.6 A Teacher ratings of inattention

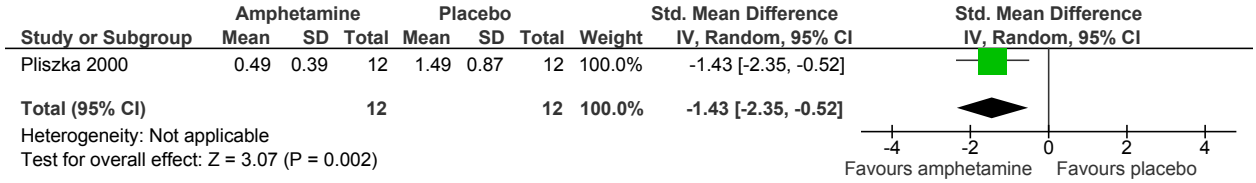


Figure 3.12.3.7: Analysis 1.7 Total ADHD symptom score-Clinician ratings

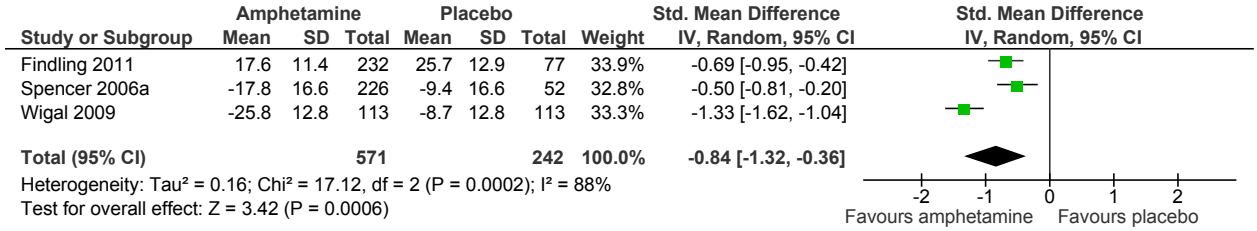


Figure 3.12.3.8: Analysis 1.8 Clinician ratings of hyperactivity/impulsivity

Study or Subgroup	Amphetamine Mean	SD	Total	Placebo Mean	SD	Total	Weight	Std. Mean Difference IV, Random, 95% CI
Findling 2011	8.5	6.56	232	11.53	6.56	77	33.9%	-0.46 [-0.72, -0.20]
Spencer 2006a	-7.6	8.6	226	-3.2	8.6	52	32.9%	-0.51 [-0.81, -0.21]
Wigal 2009	-13.3	6.8	113	-4.5	6.8	113	33.3%	-1.29 [-1.58, -1.00]
Total (95% CI)			**571**			**242**	**100.0%**	**-0.75 [-1.28, -0.23]**

Heterogeneity: Tau² = 0.20; Chi² = 20.70, df = 2 (P < 0.0001); I² = 90%
Test for overall effect: Z = 2.80 (P = 0.005)

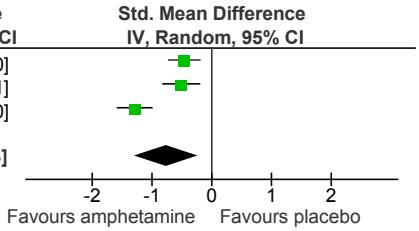


Figure 3.12.3.9: Analysis 1.9 Clinician ratings of inattention

Study or Subgroup	Amphetamine Mean	SD	Total	Placebo Mean	SD	Total	Weight	Std. Mean Difference IV, Random, 95% CI
Findling 2011	12.5	7.75	232	16.77	7.75	77	33.9%	-0.55 [-0.81, -0.29]
Spencer 2006a	-10.2	8	226	-6.1	8	52	32.8%	-0.51 [-0.82, -0.21]
Wigal 2009	-12.5	6.59	113	-4.1	6.59	113	33.3%	-1.27 [-1.56, -0.98]
Total (95% CI)			**571**			**242**	**100.0%**	**-0.78 [-1.26, -0.30]**

Heterogeneity: Tau² = 0.16; Chi² = 17.24, df = 2 (P = 0.0002); I² = 88%
Test for overall effect: Z = 3.17 (P = 0.002)

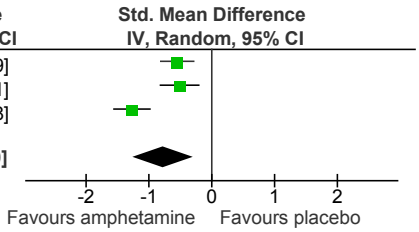


Figure 3.12.3.10: Analysis 1.10 Total ADHD symptom score-Investigator ratings

Study or Subgroup	Amphetamine Mean	SD	Total	Placebo Mean	SD	Total	Weight	Std. Mean Difference IV, Random, 95% CI
Wigal 2009	1.43	0.85	113	1.85	0.85	113	100.0%	-0.49 [-0.76, -0.23]
Total (95% CI)			**113**			**113**	**100.0%**	**-0.49 [-0.76, -0.23]**

Heterogeneity: Not applicable
Test for overall effect: Z = 3.65 (P = 0.0003)

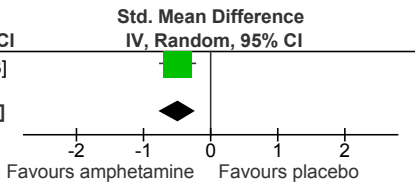


Figure 3.12.3.11: Analysis 1.11 Investigator ratings of hyperactivity/impulsivity

Study or Subgroup	Amphetamine Mean	SD	Total	Placebo Mean	SD	Total	Weight	Std. Mean Difference IV, Random, 95% CI
James 2001	2.5	1.03	35	3.8	1.1	35	100.0%	-1.21 [-1.72, -0.69]
Total (95% CI)			**35**			**35**	**100.0%**	**-1.21 [-1.72, -0.69]**

Heterogeneity: Not applicable
Test for overall effect: Z = 4.62 (P < 0.00001)

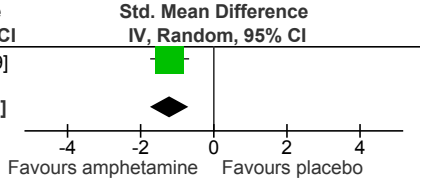


Figure 3.12.3.12: Analysis 1.12 Investigator ratings of inattention

Study or Subgroup	Amphetamine Mean	SD	Total	Placebo Mean	SD	Total	Weight	Std. Mean Difference IV, Random, 95% CI
Biederman 2007a	1.2	1	50	1.8	1	50	25.3%	-0.60 [-1.00, -0.19]
McCracken 2003	1.28	1.1	49	1.44	0.929	49	25.8%	-0.16 [-0.55, 0.24]
Wigal 2009	1.14	1.06	113	1.61	1.06	113	48.9%	-0.44 [-0.71, -0.18]
Total (95% CI)			**212**			**212**	**100.0%**	**-0.41 [-0.63, -0.19]**

Heterogeneity: Tau² = 0.01; Chi² = 2.45, df = 2 (P = 0.29); I² = 19%
Test for overall effect: Z = 3.65 (P = 0.0003)

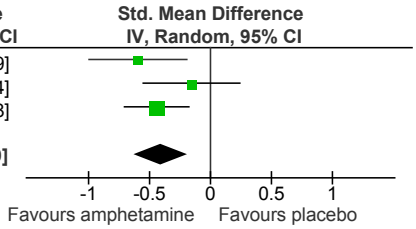


Figure 3.12.3.13: Analysis 1.13 Clinical Global Impression Severity score

Study or Subgroup	Amphetamine Mean	SD	Total	Placebo Mean	SD	Total	Weight	Std. Mean Difference IV, Random, 95% CI
Borcherding 1990	2.5	3.94	31	4.5	3.94	31	59.4%	-0.50 [-1.01, 0.00]
Pliszka 2000	1.6	0.68	12	3.22	1.44	12	40.6%	-1.39 [-2.30, -0.48]
Total (95% CI)			**43**			**43**	**100.0%**	**-0.86 [-1.72, -0.01]**

Heterogeneity: Tau² = 0.25; Chi² = 2.80, df = 1 (P = 0.09); I² = 64%
Test for overall effect: Z = 1.98 (P = 0.05)

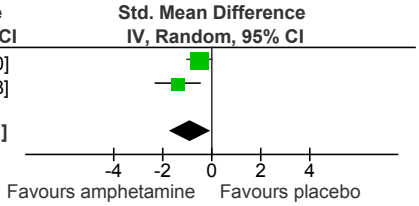


Figure 3.12.3.14: Analysis 1.14 Proportion of responders (CGI-I)

Study or Subgroup	Amphetamine Events	Total	Placebo Events	Total	Weight	Risk Ratio M-H, Random, 95% CI
Barkley 2000	16	46	5	46	6.9%	3.20 [1.28, 8.01]
Biederman 2002	148	374	35	210	16.3%	2.37 [1.71, 3.29]
Biederman 2007a	36	52	9	52	10.8%	4.00 [2.15, 7.44]
Biederman 2007b	156	218	12	72	12.5%	4.29 [2.54, 7.25]
Findling 2011	160	232	30	77	17.0%	1.77 [1.32, 2.37]
Sharp 1999	27	32	5	32	8.0%	5.40 [2.38, 12.25]
Spencer 2006a	143	233	14	63	13.4%	2.76 [1.72, 4.43]
Wigal 2009	93	129	22	129	15.0%	4.23 [2.85, 6.28]
Total (95% CI)		**1316**		**681**	**100.0%**	**3.12 [2.32, 4.20]**
Total events	779		132			

Heterogeneity: Tau² = 0.11; Chi² = 21.37, df = 7 (P = 0.003); I² = 67%
Test for overall effect: Z = 7.51 (P < 0.00001)

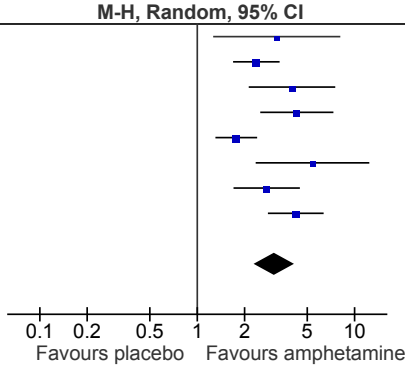


Figure 3.12.3.15: Analysis 1.15 Quality of life

Study or Subgroup	Amphetamine Mean	SD	Total	Placebo Mean	SD	Total	Weight	Std. Mean Difference IV, Random, 95% CI
Findling 2011	81.2	12.53	232	81.3	12.16	77	100.0%	-0.01 [-0.27, 0.25]
Total (95% CI)			**232**			**77**	**100.0%**	**-0.01 [-0.27, 0.25]**

Heterogeneity: Not applicable
Test for overall effect: Z = 0.06 (P = 0.95)

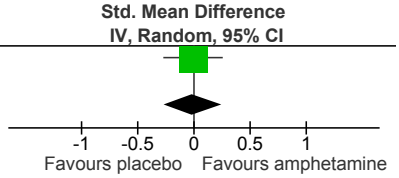


Figure 3.12.3.16: Analysis 1.16 Academic performance

Study or Subgroup	Amphetamine Mean	SD	Total	Placebo Mean	SD	Total	Weight	Std. Mean Difference IV, Random, 95% CI
Biederman 2007a	129.5	76	50	84.1	76	50	16.5%	0.59 [0.19, 0.99]
Borcherding 1990	97.1	4.6	33	94	7.9	33	12.3%	0.47 [-0.02, 0.96]
James 2001	169.9	52.7	35	140.2	51.3	35	12.8%	0.56 [0.09, 1.04]
McCracken 2003	79.11	52.03	49	64.88	50.3	49	16.7%	0.28 [-0.12, 0.67]
Nemzer 1986	45.78	16.7	14	37.78	17.4	14	6.0%	0.46 [-0.30, 1.21]
Shekim 1986	95.06	9.46	22	95.12	11.67	22	9.1%	-0.01 [-0.60, 0.59]
Wigal 2009	109.23	36.28	113	80.82	36.28	113	26.6%	0.78 [0.51, 1.05]
Total (95% CI)			**316**			**316**	**100.0%**	**0.51 [0.31, 0.70]**

Heterogeneity: Tau² = 0.02; Chi² = 8.19, df = 6 (P = 0.22); I² = 27%
Test for overall effect: Z = 5.11 (P < 0.00001)

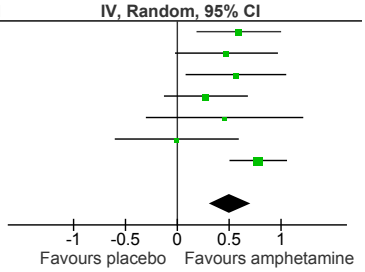


Figure 3.12.3.17: Analysis 1.17 Proportion of participants who completed the trial

Study or Subgroup	Amphetamine Events	Total	Placebo Events	Total	Weight	Risk Ratio M-H, Random, 95% CI
Biederman 2002	298	374	136	210	12.1%	1.23 [1.10, 1.38]
Biederman 2007a	52	52	50	52	16.9%	1.04 [0.97, 1.11]
Biederman 2007b	176	218	54	72	9.2%	1.08 [0.93, 1.25]
Findling 2011	194	233	69	79	13.1%	0.95 [0.86, 1.06]
Pliszka 2000	18	20	16	18	5.5%	1.01 [0.81, 1.26]
Sharp 1999	32	32	31	32	14.7%	1.03 [0.95, 1.12]
Spencer 2006a	187	233	48	63	8.9%	1.05 [0.90, 1.23]
Wigal 2009	127	129	125	129	19.5%	1.02 [0.98, 1.06]
Total (95% CI)		**1291**		**655**	**100.0%**	**1.05 [0.99, 1.11]**
Total events	1084		529			

Heterogeneity: Tau² = 0.00; Chi² = 22.66, df = 7 (P = 0.002); I² = 69%
Test for overall effect: Z = 1.48 (P = 0.14)

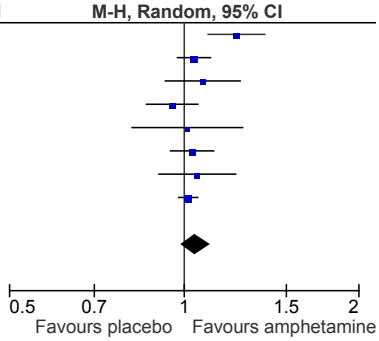


Figure 3.12.3.18: Analysis 1.18 Proportion of participants who experienced at least one adverse event

Study or Subgroup	Favours amphetamine Events	Total	Placebo Events	Total	Weight	Risk Ratio M-H, Random, 95% CI
Biederman 2002	263	374	119	210	47.3%	1.24 [1.08, 1.42]
Biederman 2007a	9	52	8	52	1.9%	1.13 [0.47, 2.69]
Biederman 2007b	162	218	34	72	18.7%	1.57 [1.22, 2.03]
Findling 2011	160	233	45	79	25.7%	1.21 [0.98, 1.49]
Wigal 2009	38	129	22	129	6.4%	1.73 [1.09, 2.75]
Total (95% CI)		**1006**		**542**	**100.0%**	**1.31 [1.16, 1.48]**
Total events	632		228			

Heterogeneity: Tau² = 0.00; Chi² = 4.76, df = 4 (P = 0.31); I² = 16%
Test for overall effect: Z = 4.41 (P < 0.0001)

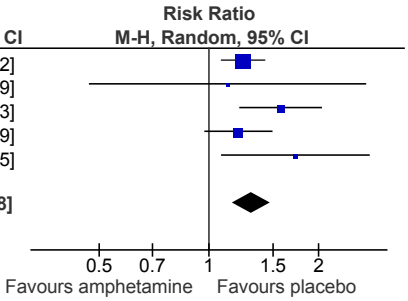


Figure 3.12.3.19: Analysis 1.19 Proportion of participants who withdraw due to an adverse event

Study or Subgroup	Amphetamine Events	Amphetamine Total	Placebo Events	Placebo Total	Weight	Risk Ratio M-H, Random, 95% CI
Biederman 2002	9	374	6	210	48.2%	0.84 [0.30, 2.33]
Biederman 2007b	21	218	1	72	12.7%	6.94 [0.95, 50.65]
Borcherding 1990	1	46	0	46	5.0%	3.00 [0.13, 71.78]
Findling 2011	10	233	1	79	12.0%	3.39 [0.44, 26.07]
Pliszka 2000	2	20	0	18	5.7%	4.52 [0.23, 88.38]
Spencer 2006a	5	233	0	63	6.0%	3.01 [0.17, 53.69]
Swanson 1998a	2	33	0	33	5.6%	5.00 [0.25, 100.32]
Wigal 2009	0	129	1	129	4.9%	0.33 [0.01, 8.11]
Total (95% CI)		**1286**		**650**	**100.0%**	**1.74 [0.86, 3.52]**
Total events	50		9			

Heterogeneity: Tau² = 0.00; Chi² = 6.73, df = 7 (P = 0.46); I² = 0%
Test for overall effect: Z = 1.53 (P = 0.13)

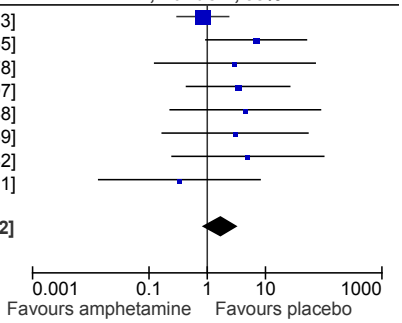


Figure 3.12.3.20: Analysis 1.20 Proportion of participants who experience decreased appetite

Study or Subgroup	Amphetamine Events	Amphetamine Total	Placebo Events	Placebo Total	Weight	Risk Ratio M-H, Random, 95% CI
Biederman 2002	82	374	4	210	16.9%	11.51 [4.28, 30.96]
Biederman 2007a	2	52	0	52	6.4%	5.00 [0.25, 101.68]
Biederman 2007b	85	218	3	72	16.0%	9.36 [3.05, 28.68]
Findling 2011	79	233	2	79	14.2%	13.39 [3.37, 53.23]
McCracken 2003	20	51	11	51	19.2%	1.82 [0.97, 3.40]
Pliszka 2000	3	20	0	18	6.7%	6.33 [0.35, 114.81]
Spencer 2006a	83	233	1	63	10.7%	22.44 [3.19, 158.05]
Wigal 2009	7	129	1	129	10.0%	7.00 [0.87, 56.09]
Total (95% CI)		**1310**		**674**	**100.0%**	**7.44 [2.99, 18.48]**
Total events	361		22			

Heterogeneity: Tau² = 1.02; Chi² = 23.09, df = 7 (P = 0.002); I² = 70%
Test for overall effect: Z = 4.32 (P < 0.0001)

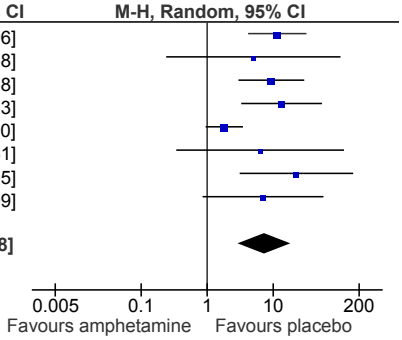


Figure 3.12.3.21: Analysis 1.21 Proportion of participants who experience insomnia/sleep problems

Study or Subgroup	Amphetamines Events	Amphetamines Total	Placebo Events	Placebo Total	Weight	Risk Ratio M-H, Random, 95% CI
Biederman 2002	62	374	4	210	19.6%	8.70 [3.21, 23.58]
Biederman 2007a	1	52	1	52	5.4%	1.00 [0.06, 15.57]
Biederman 2007b	41	218	2	72	14.2%	6.77 [1.68, 27.29]
Findling 2011	26	233	3	79	17.1%	2.94 [0.91, 9.44]
McCracken 2003	16	51	10	51	24.8%	1.60 [0.80, 3.18]
Spencer 2006a	28	233	2	63	14.0%	3.79 [0.93, 15.46]
Wigal 2009	5	129	0	129	5.0%	11.00 [0.61, 196.91]
Total (95% CI)		**1290**		**656**	**100.0%**	**3.67 [1.83, 7.38]**
Total events	179		22			

Heterogeneity: Tau² = 0.39; Chi² = 11.73, df = 6 (P = 0.07); I² = 49%
Test for overall effect: Z = 3.65 (P = 0.0003)

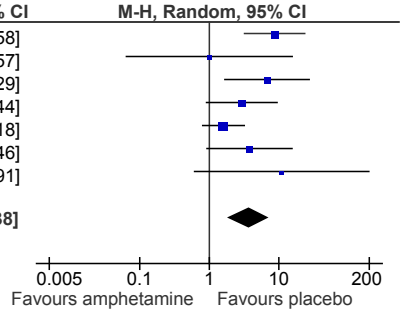


Figure 3.12.3.22: Analysis 1.22 Proportion of participants who experience abdominal pain

Study or Subgroup	Amphetamines Events	Total	Placebo Events	Total	Weight	Risk Ratio M-H, Random, 95% CI
Biederman 2002	54	374	20	210	48.9%	1.52 [0.93, 2.46]
Biederman 2007a	2	52	1	52	2.0%	2.00 [0.19, 21.38]
Biederman 2007b	26	218	4	72	11.1%	2.15 [0.78, 5.94]
McCracken 2003	18	51	12	51	30.0%	1.50 [0.81, 2.78]
Pliszka 2000	5	20	0	18	1.4%	9.95 [0.59, 168.27]
Spencer 2006a	25	233	1	63	2.9%	6.76 [0.93, 48.92]
Wigal 2009	2	129	3	129	3.7%	0.67 [0.11, 3.92]
Total (95% CI)		**1077**		**595**	**100.0%**	**1.65 [1.17, 2.31]**
Total events	132		41			

Heterogeneity: Tau² = 0.00; Chi² = 5.27, df = 6 (P = 0.51); I² = 0%
Test for overall effect: Z = 2.88 (P = 0.004)

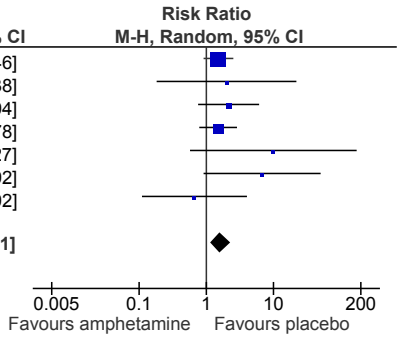


Figure 3.12.3.23: Analysis 1.23 Proportion of participants who experience headaches

Study or Subgroup	Amphetamine Events	Total	Placebo Events	Total	Weight	Risk Ratio M-H, Random, 95% CI
Biederman 2002	67	374	45	210	53.3%	0.84 [0.60, 1.17]
Biederman 2007b	26	218	7	72	9.7%	1.23 [0.56, 2.71]
Findling 2011	34	233	10	79	14.1%	1.15 [0.60, 2.22]
Giblin 2011	5	16	1	8	1.6%	2.50 [0.35, 17.97]
Pliszka 2000	2	20	1	18	1.1%	1.80 [0.18, 18.21]
Spencer 2006a	38	233	12	63	17.7%	0.86 [0.48, 1.54]
Wigal 2009	6	129	2	129	2.4%	3.00 [0.62, 14.59]
Total (95% CI)		**1223**		**579**	**100.0%**	**0.97 [0.75, 1.24]**
Total events	178		78			

Heterogeneity: Tau² = 0.00; Chi² = 4.69, df = 6 (P = 0.58); I² = 0%
Test for overall effect: Z = 0.28 (P = 0.78)

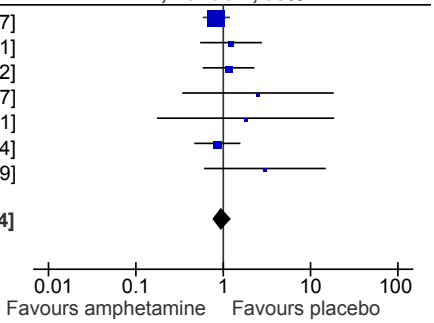


Figure 3.12.3.24: Analysis 1.24 Proportion of participants who experience anxiety

Study or Subgroup	Amphetamine Events	Total	Placebo Events	Total	Weight	Risk Ratio M-H, Random, 95% CI
Biederman 2002	21	374	4	210	23.9%	2.95 [1.03, 8.47]
McCracken 2003	39	51	39	51	49.7%	1.00 [0.81, 1.24]
Pliszka 2000	1	20	1	18	6.0%	0.90 [0.06, 13.36]
Spencer 2006a	14	233	3	63	20.4%	1.26 [0.37, 4.25]
Total (95% CI)		**678**		**342**	**100.0%**	**1.35 [0.67, 2.73]**
Total events	75		47			

Heterogeneity: Tau² = 0.25; Chi² = 6.07, df = 3 (P = 0.11); I² = 51%
Test for overall effect: Z = 0.84 (P = 0.40)

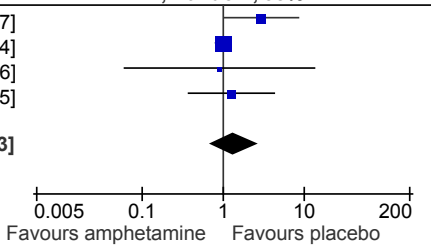
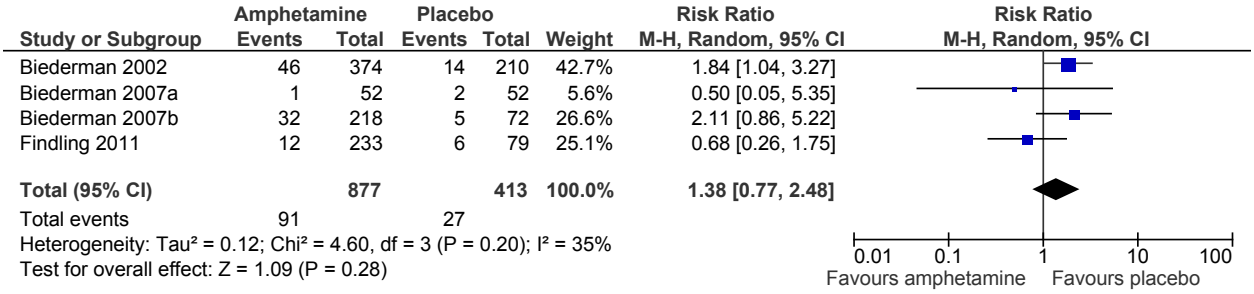


Figure 3.12.3.25: Analysis 1.25 Proportion of participants who experience nausea



3.12.4 Subgroup Analysis 1: Type of amphetamine

Figure 3.12.4.1: Subgroup Analysis 2.1 Total ADHD symptom score-Parent ratings

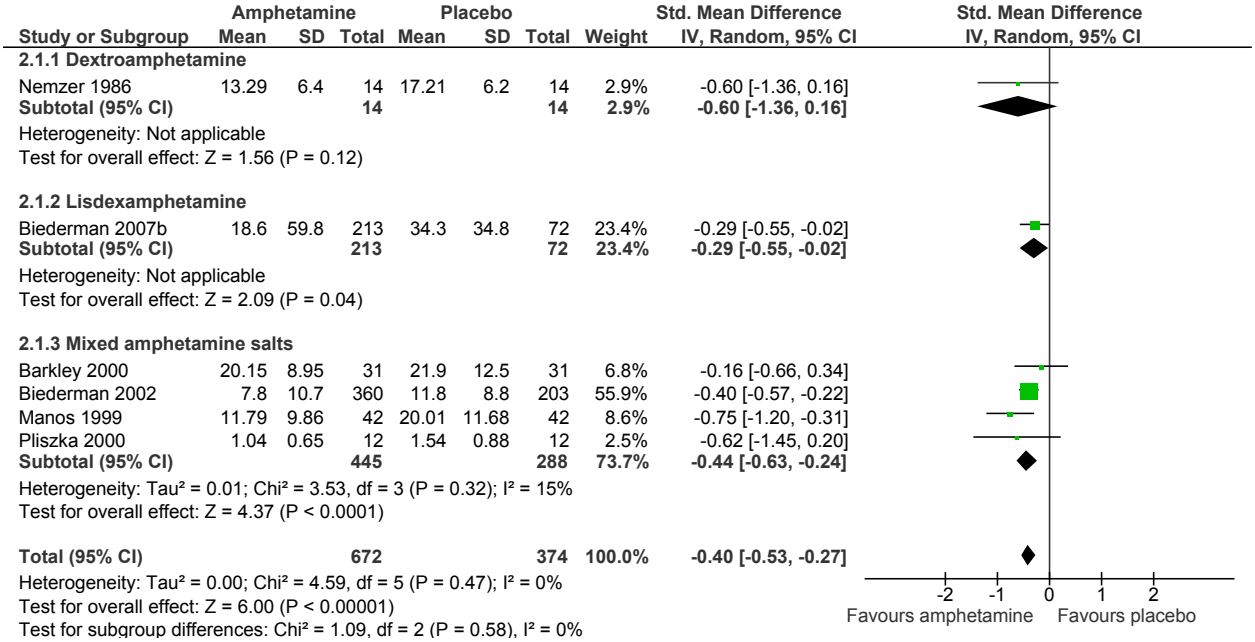


Figure 3.12.4.2: Subgroup Analysis 2.2 Total ADHD symptom score-Clinician ratings

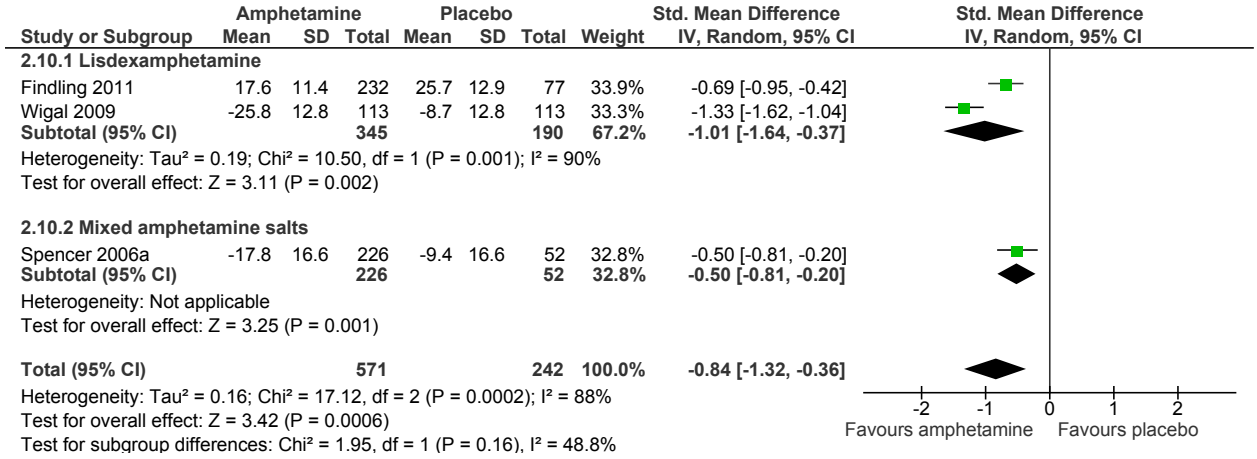


Figure 3.12.4.3: Subgroup Analysis 2.3 Clinician ratings of hyperactivity/impulsivity

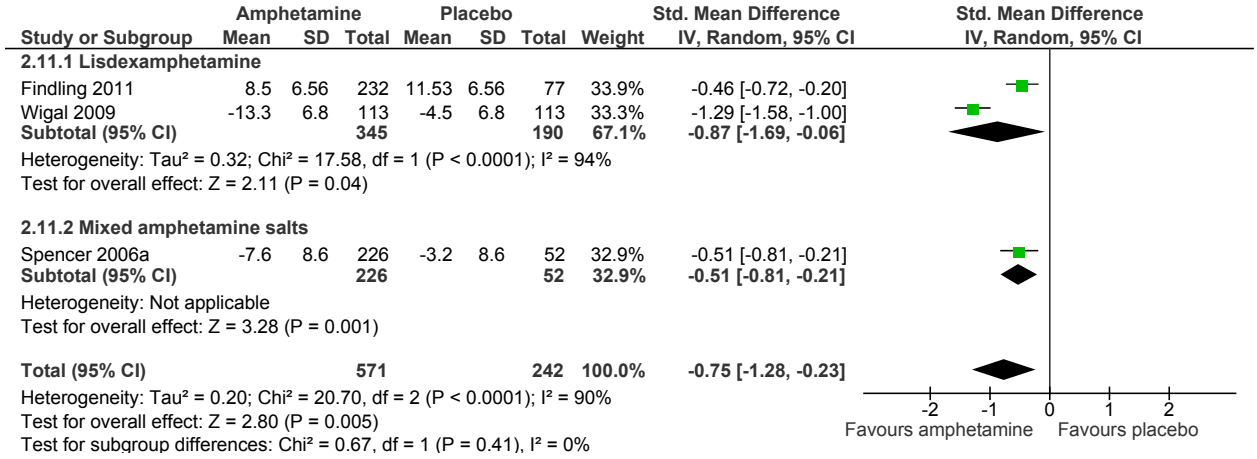


Figure 3.12.4.4: Subgroup Analysis 2.4 Clinician ratings of inattention

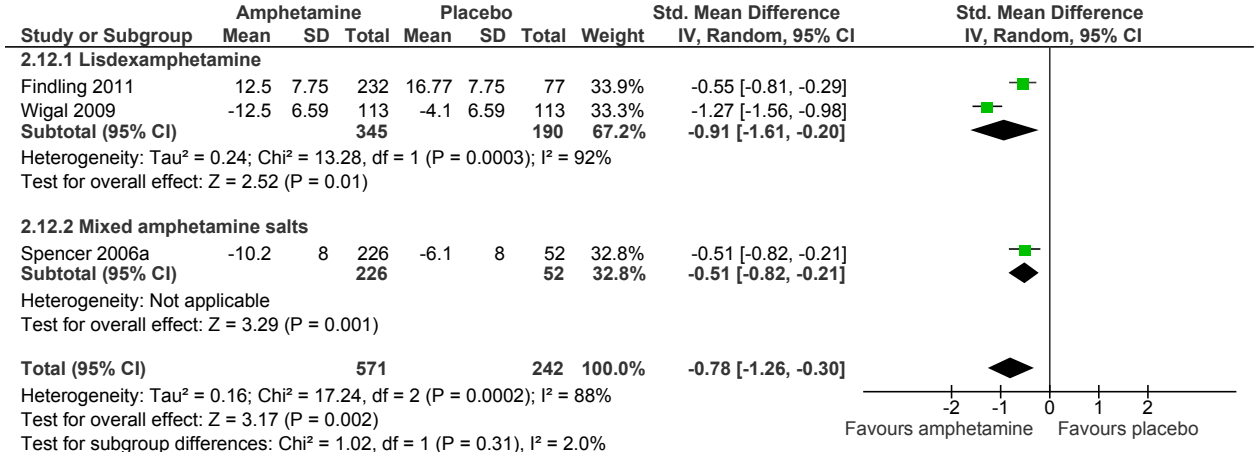


Figure 3.12.4.5: Subgroup Analysis 2.5 Proportion of responders (CGI-I)

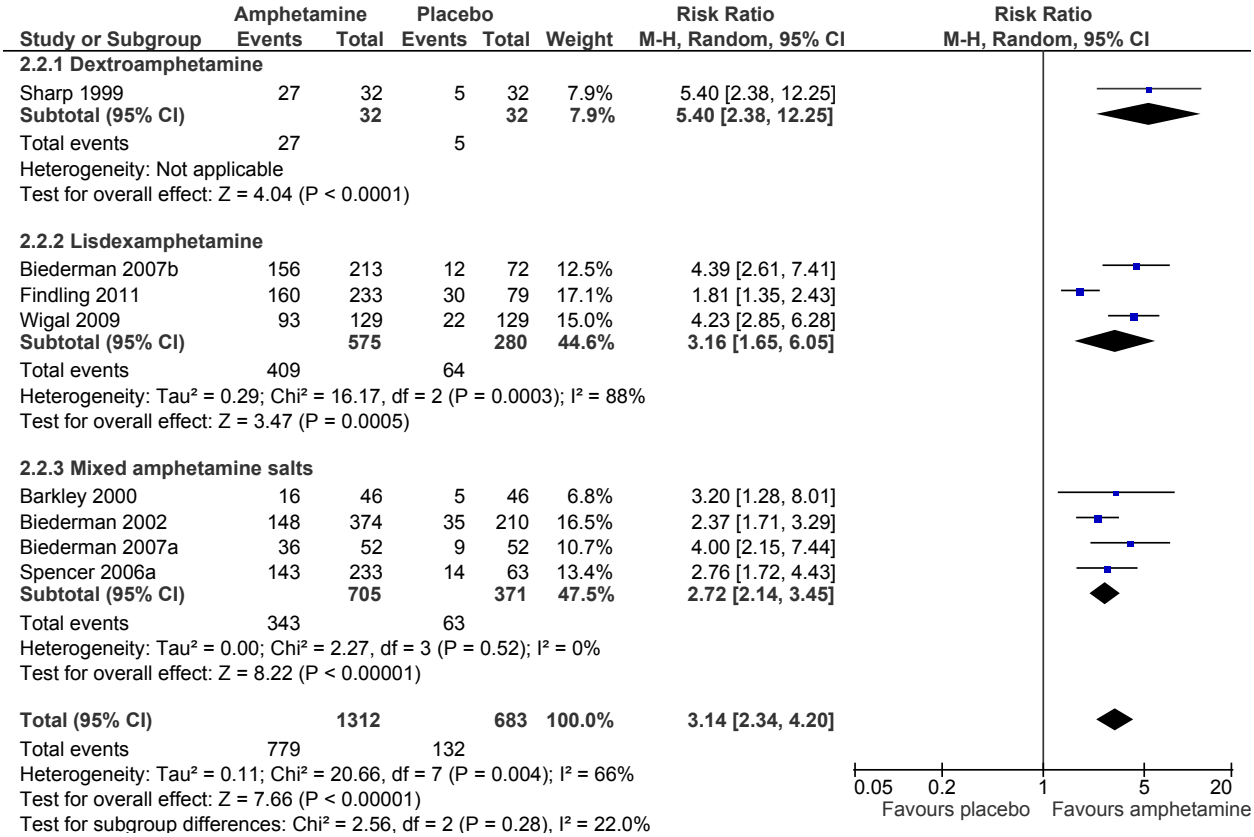


Figure 3.12.4.6: Subgroup Analysis 2.6 Academic performance

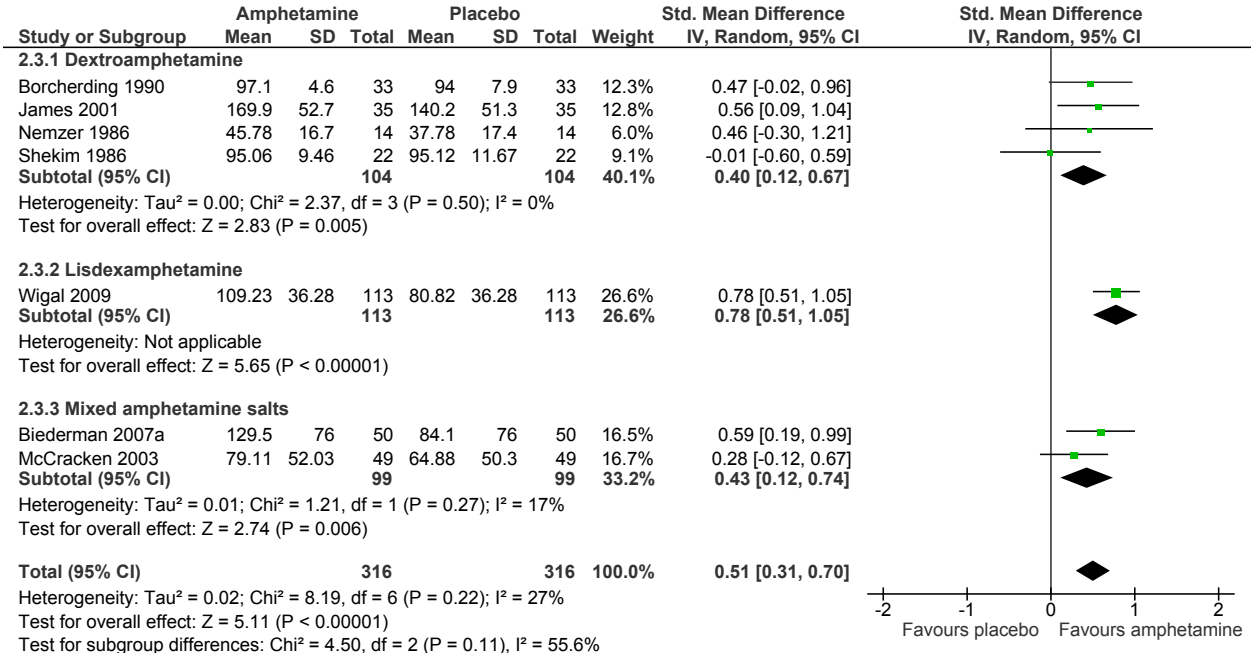


Figure 3.12.4.7: Subgroup Analysis 2.7 Proportion of participants who completed the trial

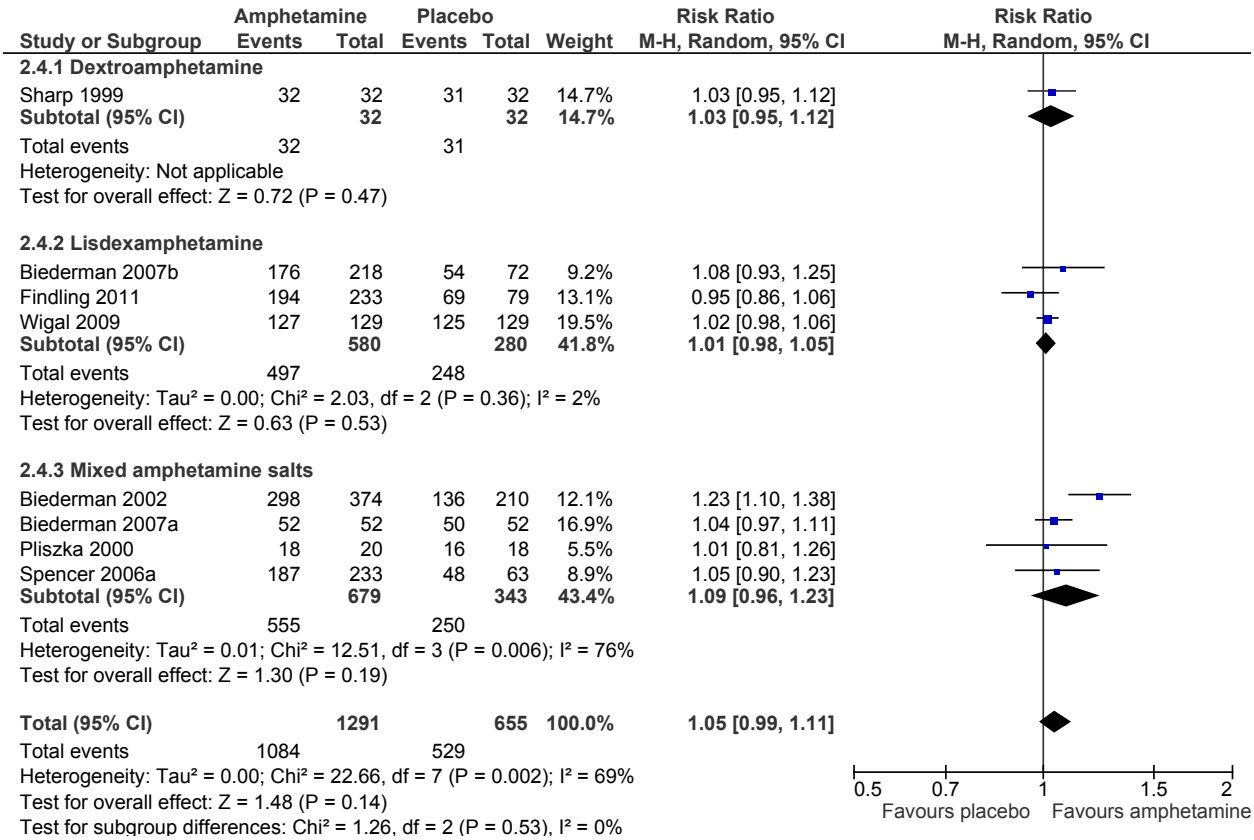


Figure 3.12.4.8: Subgroup Analysis 2.8 Proportion of participants who withdrew due to an adverse event

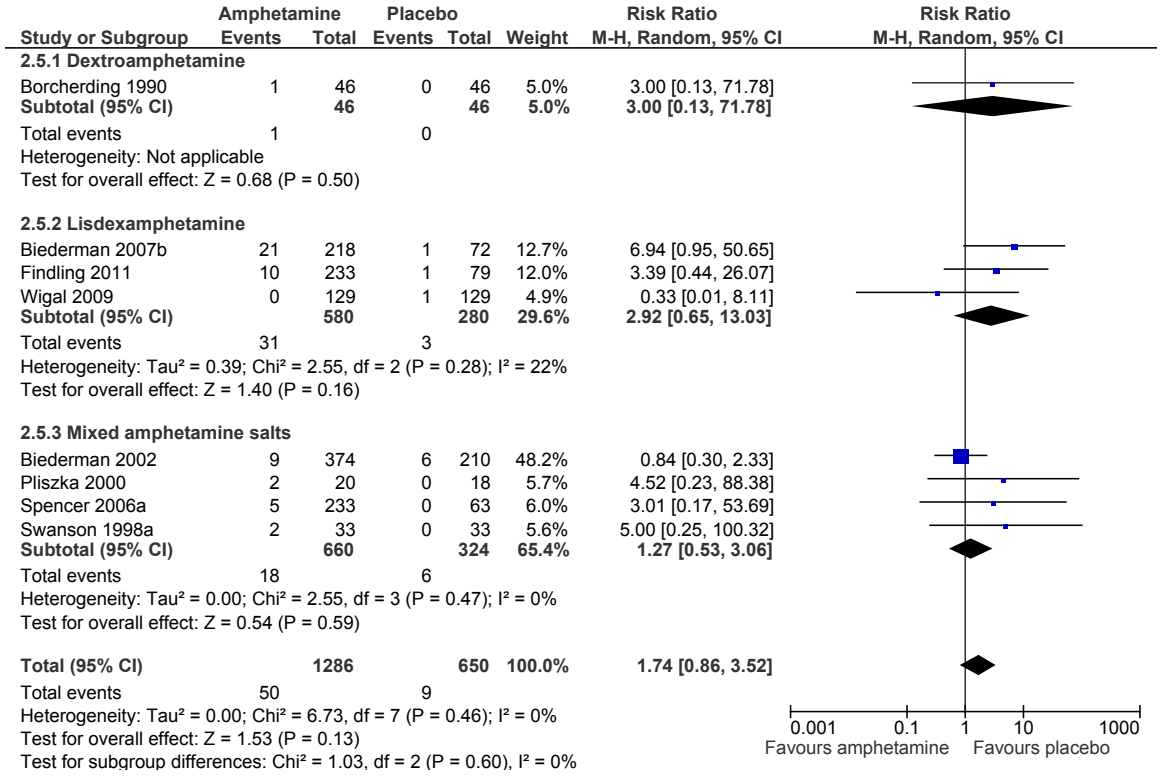


Figure 3.12.4.9: Subgroup Analysis 2.9 Proportion of participants who experience decreased appetite

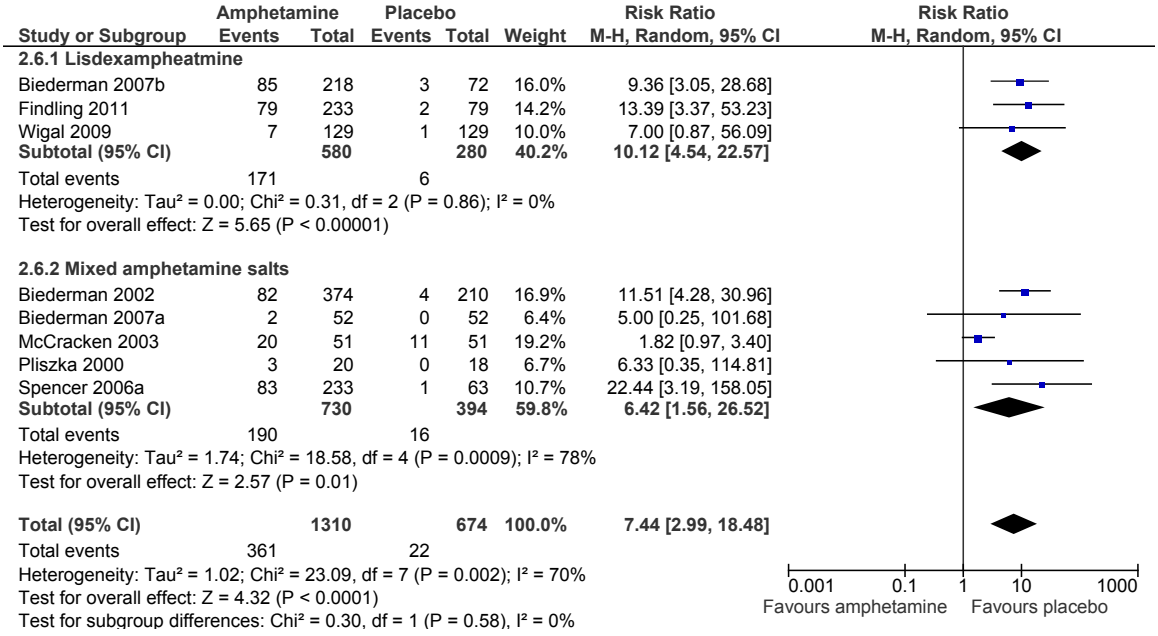


Figure 3.12.4.10: Subgroup Analysis 2.10 Proportion of participants who experience insomnia/sleep problems

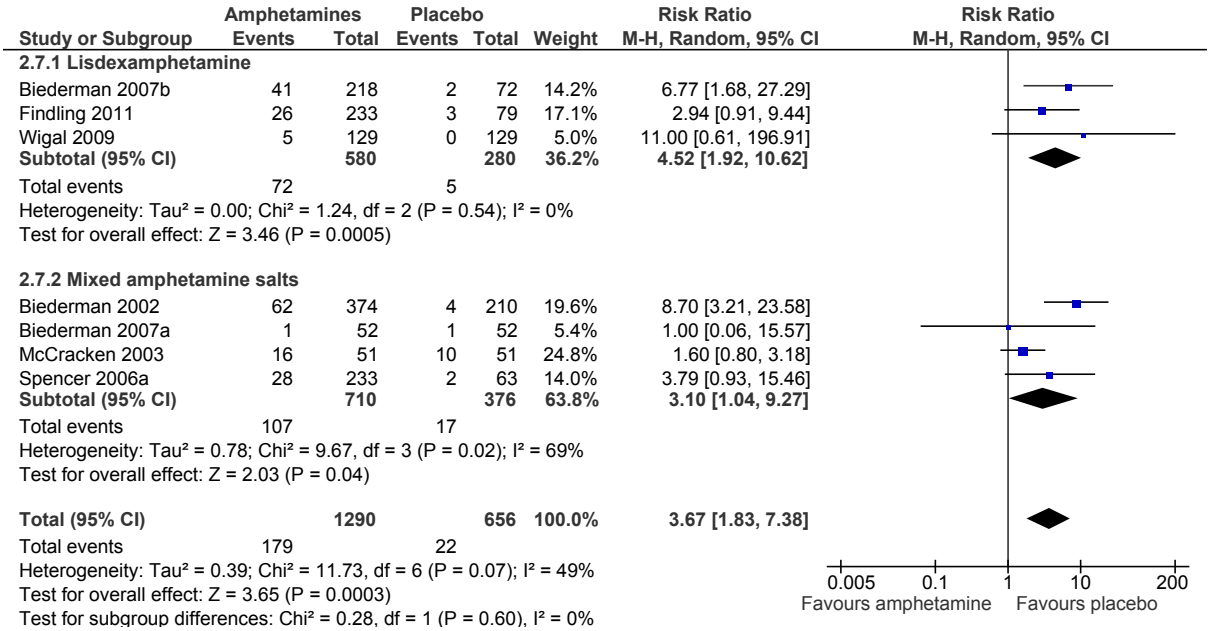


Figure 3.12.4.11: Subgroup Analysis 2.11 Proportion of participants who experience abdominal pain

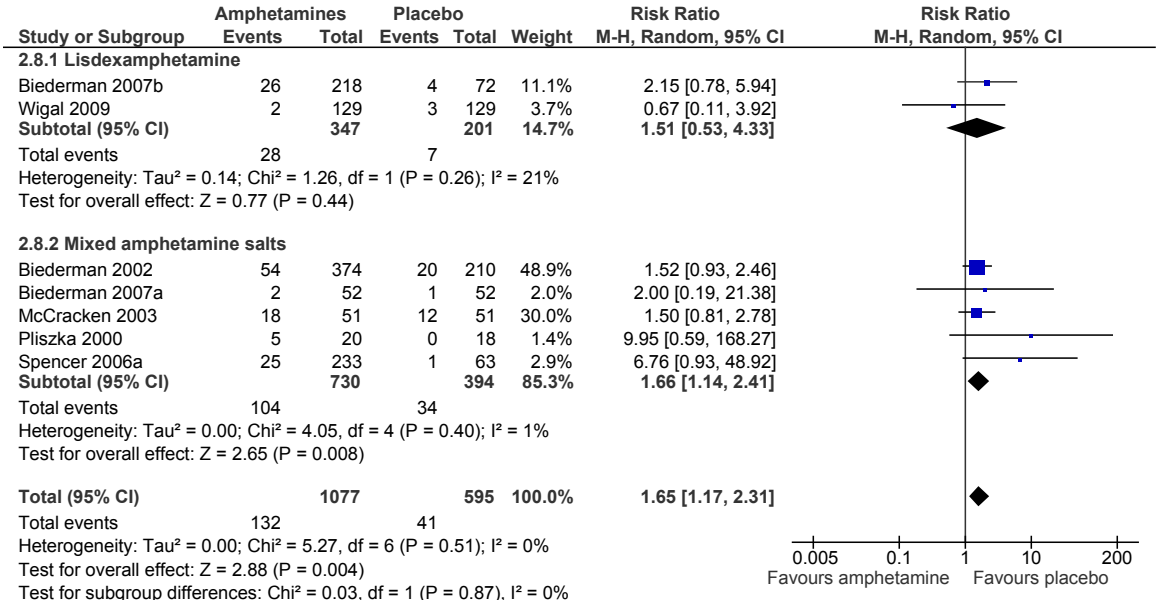
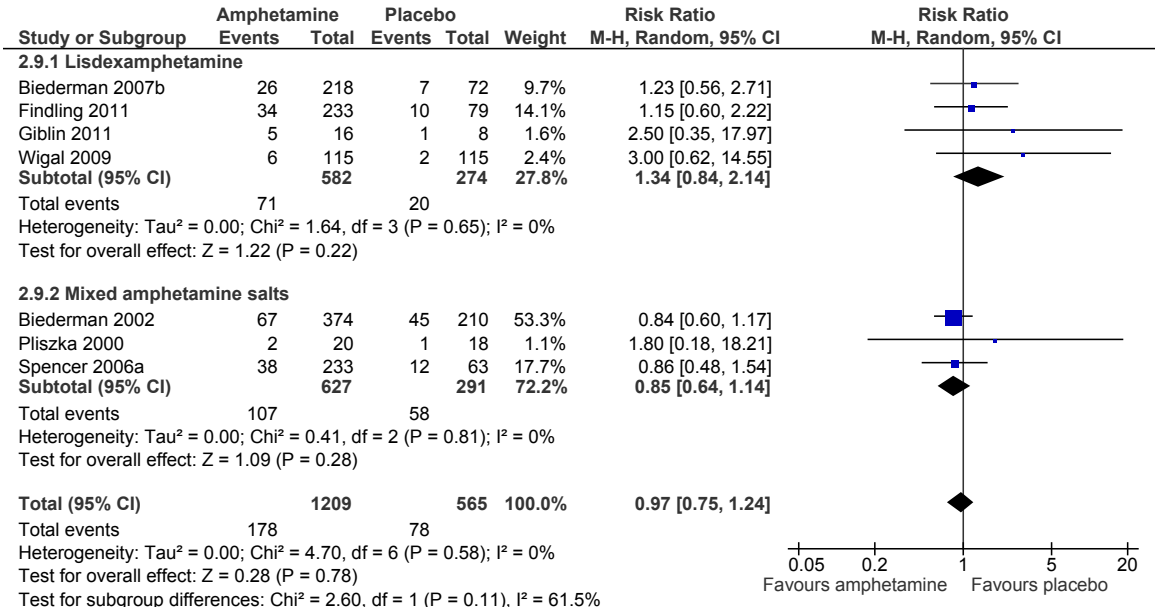


Figure 3.12.4.12: Subgroup Analysis 2.12 Proportion of participants who experience headaches



3.12.5 Subgroup Analysis 2: Amphetamine release formulation

Figure 3.12.5.1: Subgroup Analysis 3.1 Total ADHD symptom score-Parent ratings

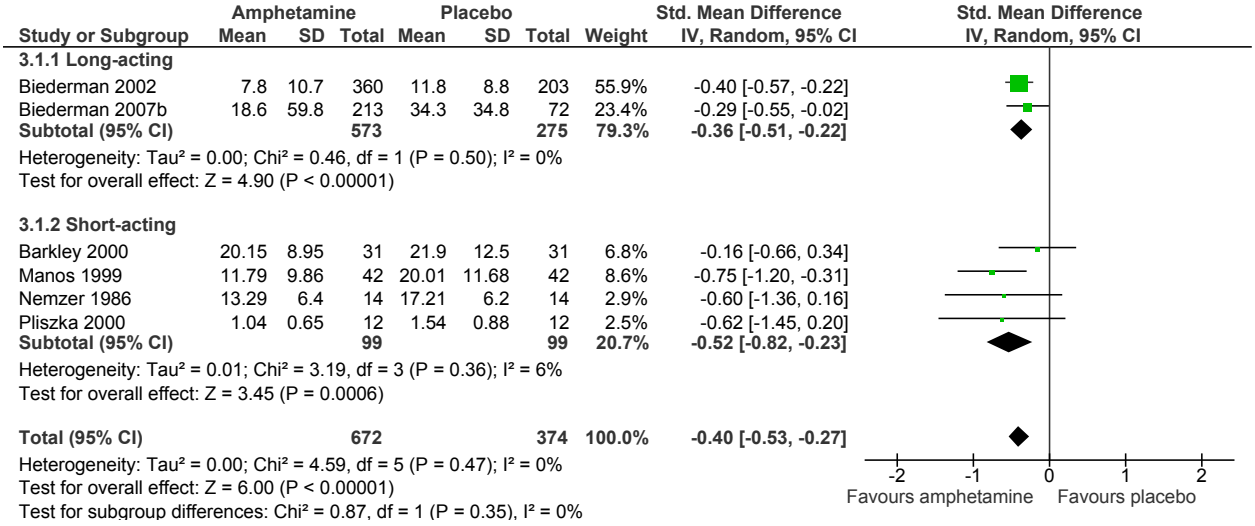


Figure 3.12.5.2: Subgroup Analysis 3.2 Proportion of responders (CGI-I)

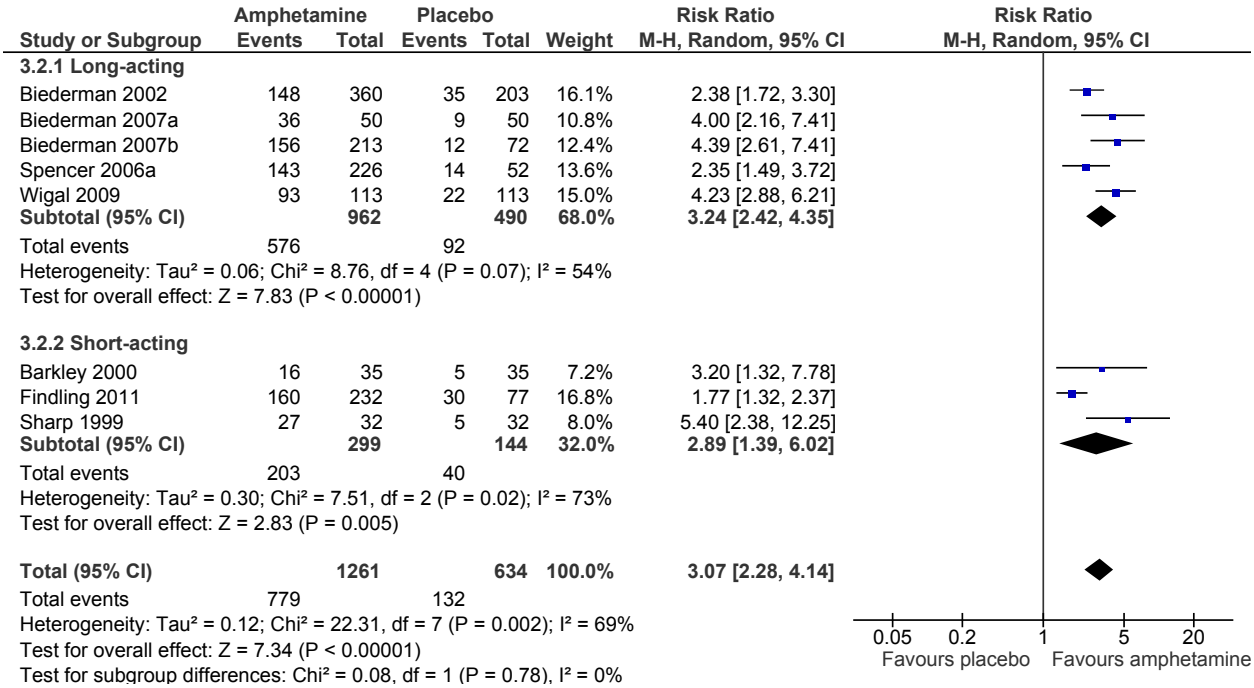


Figure 3.12.5.3: Subgroup Analysis 3.3 Academic performance

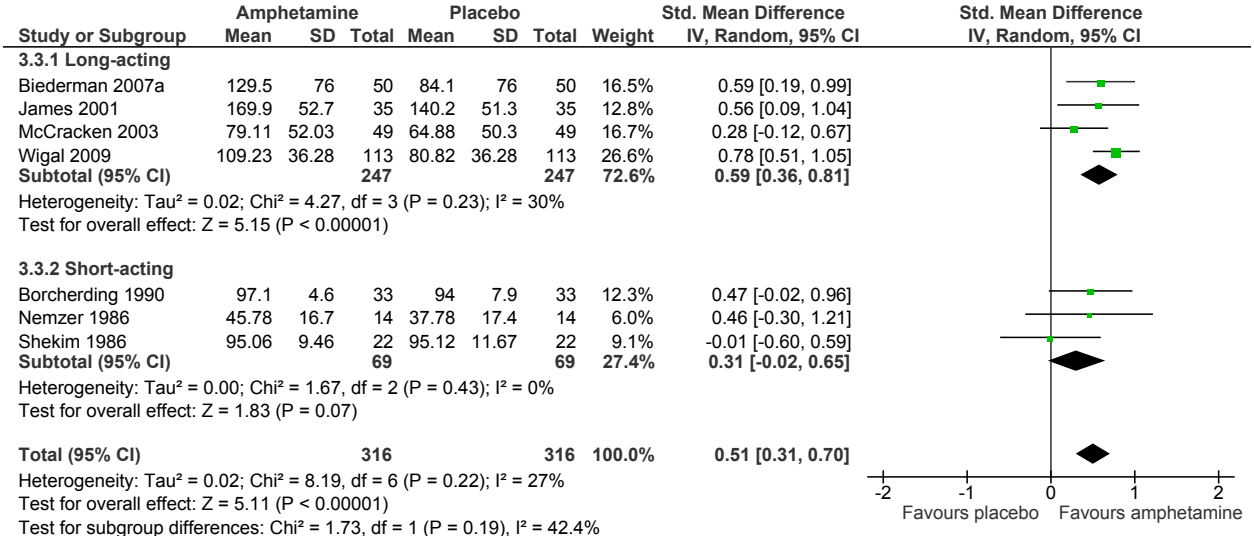


Figure 3.12.5.4: Subgroup Analysis 3.4 Proportion of participants who completed the trial

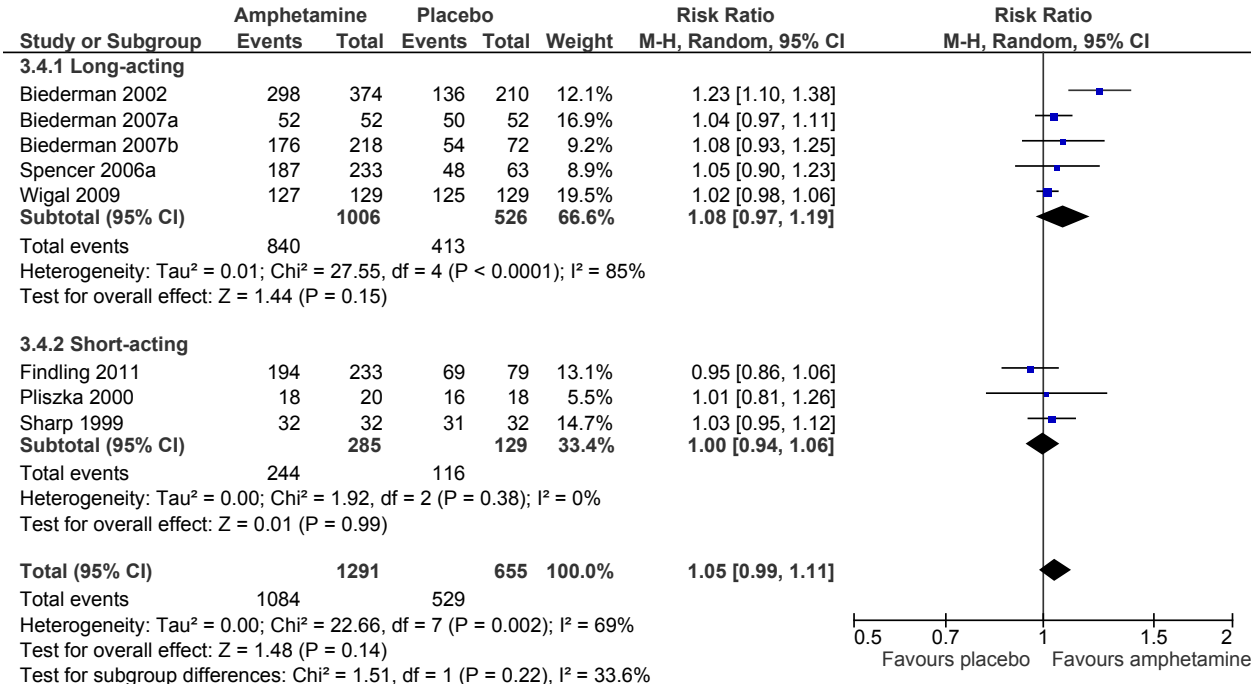


Figure 3.12.5.5: Subgroup Analysis 3.5 Proportion of participants who experience decreased appetite

Study or Subgroup	Amphetamine Events	Total	Placebo Events	Total	Weight	Risk Ratio M-H, Random, 95% CI
3.5.1 Long-acting						
Biederman 2002	82	374	4	210	16.9%	11.51 [4.28, 30.96]
Biederman 2007a	2	52	0	52	6.3%	5.00 [0.25, 101.68]
Biederman 2007b	85	218	3	72	16.0%	9.36 [3.05, 28.68]
Findling 2011	79	233	2	77	14.2%	13.05 [3.29, 51.86]
McCracken 2003	20	51	11	51	19.2%	1.82 [0.97, 3.40]
Spencer 2006a	83	233	1	63	10.7%	22.44 [3.19, 158.05]
Wigal 2009	7	129	1	129	10.0%	7.00 [0.87, 56.09]
Subtotal (95% CI)		**1290**		**654**	**93.3%**	**7.54 [2.85, 19.95]**
Total events	358		22			
Heterogeneity: Tau² = 1.12; Chi² = 23.00, df = 6 (P = 0.0008); I² = 74%						
Test for overall effect: Z = 4.07 (P < 0.0001)						
3.5.2 Short-acting						
Pliszka 2000	3	20	0	18	6.7%	6.33 [0.35, 114.81]
Subtotal (95% CI)		**20**		**18**	**6.7%**	**6.33 [0.35, 114.81]**
Total events	3		0			
Heterogeneity: Not applicable						
Test for overall effect: Z = 1.25 (P = 0.21)						
Total (95% CI)		**1310**		**672**	**100.0%**	**7.41 [2.99, 18.36]**
Total events	361		22			
Heterogeneity: Tau² = 1.02; Chi² = 22.95, df = 7 (P = 0.002); I² = 69%						
Test for overall effect: Z = 4.32 (P < 0.0001)						
Test for subgroup differences: Chi² = 0.01, df = 1 (P = 0.91), I² = 0%						

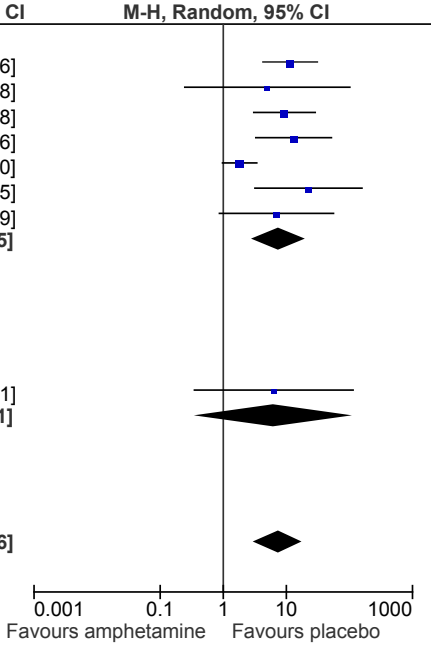
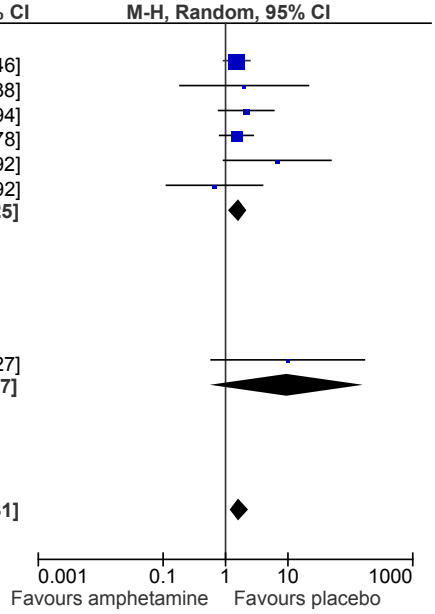


Figure 3.12.5.6: Subgroup Analysis 3.6 Proportion of participants who experience abdominal pain

Study or Subgroup	Amphetamines Events	Total	Placebo Events	Total	Weight	Risk Ratio M-H, Random, 95% CI
3.6.1 Long-acting						
Biederman 2002	54	374	20	210	48.9%	1.52 [0.93, 2.46]
Biederman 2007a	2	52	1	52	2.0%	2.00 [0.19, 21.38]
Biederman 2007b	26	218	4	72	11.1%	2.15 [0.78, 5.94]
McCracken 2003	18	51	12	51	30.0%	1.50 [0.81, 2.78]
Spencer 2006a	25	233	1	63	2.9%	6.76 [0.93, 48.92]
Wigal 2009	2	129	3	129	3.7%	0.67 [0.11, 3.92]
Subtotal (95% CI)		**1057**		**577**	**98.6%**	**1.60 [1.14, 2.25]**
Total events	127		41			
Heterogeneity: Tau² = 0.00; Chi² = 3.57, df = 5 (P = 0.61); I² = 0%						
Test for overall effect: Z = 2.71 (P = 0.007)						
3.6.2 Short-acting						
Pliszka 2000	5	20	0	18	1.4%	9.95 [0.59, 168.27]
Subtotal (95% CI)		**20**		**18**	**1.4%**	**9.95 [0.59, 168.27]**
Total events	5		0			
Heterogeneity: Not applicable						
Test for overall effect: Z = 1.59 (P = 0.11)						
Total (95% CI)		**1077**		**595**	**100.0%**	**1.65 [1.17, 2.31]**
Total events	132		41			
Heterogeneity: Tau² = 0.00; Chi² = 5.27, df = 6 (P = 0.51); I² = 0%						
Test for overall effect: Z = 2.88 (P = 0.004)						
Test for subgroup differences: Chi² = 1.58, df = 1 (P = 0.21), I² = 36.7%						



3.12.6 Subgroup Analysis 3: Funding source

Figure 3.12.6.1: Subgroup Analysis 4.1 Total ADHD symptom score-Parent ratings

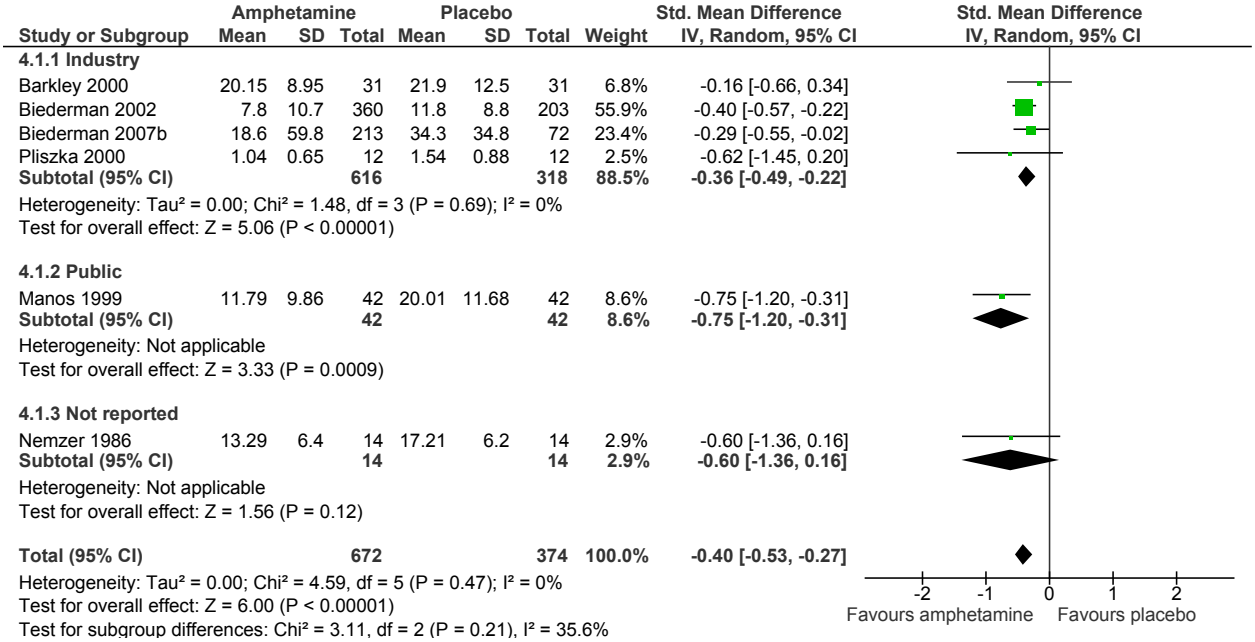


Figure 3.12.6.2: Subgroup Analysis 4.2 Proportion of responders (CGI-I)

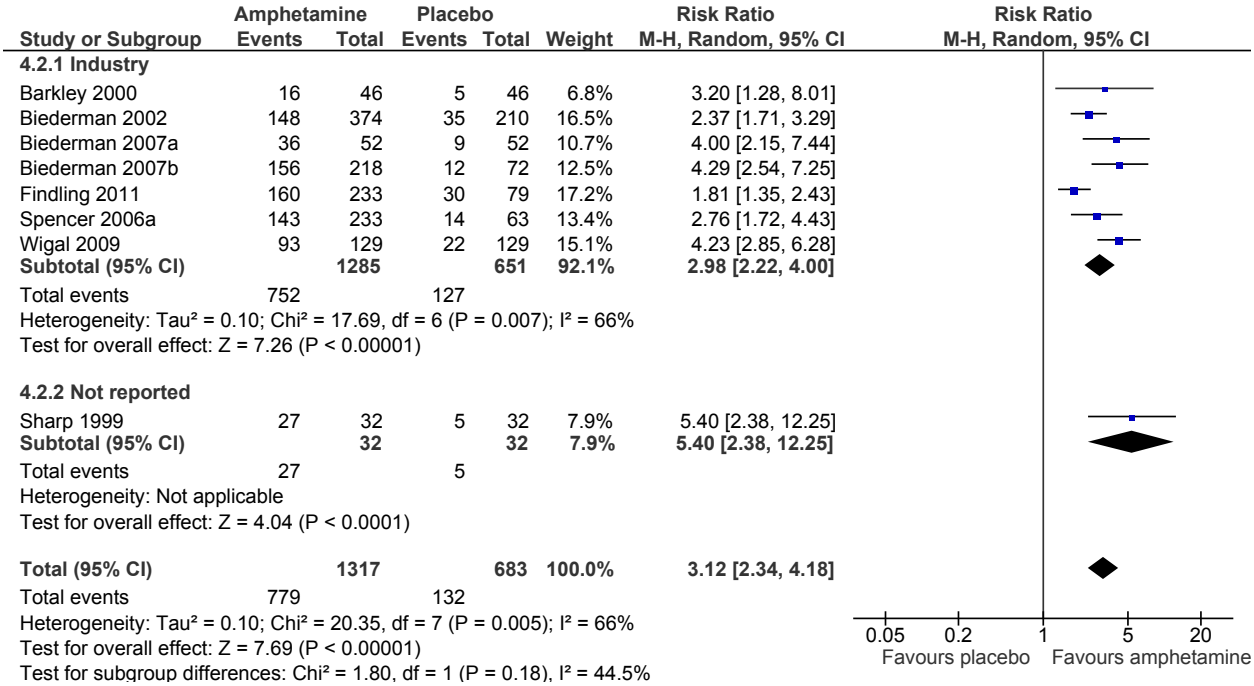
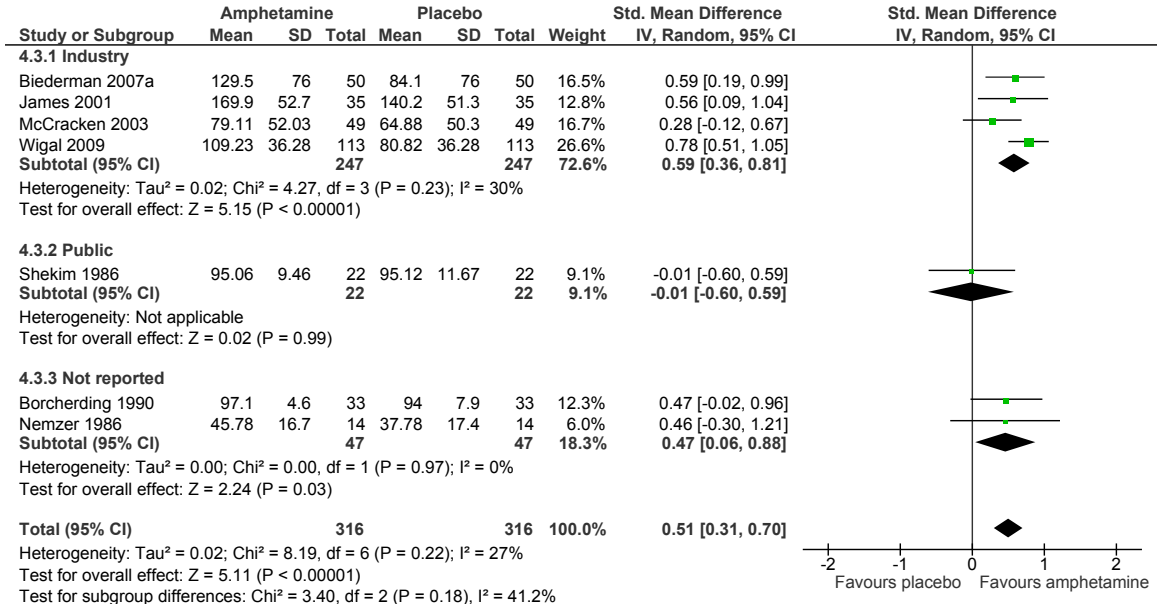


Figure 3.12.6.3: Subgroup Analysis 4.3 Academic performance



3.13 APPENDICES

3.13.1 MEDLINE Search Strategy
Ovid platform; 1948-July 9, 2013

1. exp Amphetamines/
2. (amphetamine$ or dexamphetamine$ or methamphetamine$ or dextroamphetamine$ or lisdexamphetamine$ or vyvanase$ or Dexedrin3 or desoxyn$ or adderall$).mp.
3. Central Nervous System Stimulants/
4. 1 or 2 or 3
5. exp Attention Deficit Disorder with Hyperactivity/
6. Child Behavior Disorders/
7. adhd.tw.
8. addh.tw.
9. adhs.tw.
10. adhs.tw.
11. "ad/hd".tw.
12. hyperactiv$.tw.
13. hyper-activ$.tw.
14. overactiv$.tw.
15. over-activ$.tw.
16. hyperkinesis/
17. hyperkin$.tw.
18. hyper-kin$.tw.
19. hkd.tw.
20. (minimal adj3 brain$ adj3 (damag$ or disorder$ or dysfunc$)).tw.
21. (attention$ adj3 (deficit$ or disorder$ or dysfunc$)).tw.
22. (behav$ adj3 (dysfunc$ or disorder$)).tw.
23. (behav$ adj3 (dysfunc$ or disorder$)).tw.
24. (impulsiv$ or inattentiv$ or inattention$).tw.
25. disruptiv$.tw.
26. or/5-25
27. exp child/
28. adolescent/
29. (adoles$ or teen$ or youth$ or young people or young person$).tw.
30. (child$ or toddler$ or preschool$ or pre-school or schoolchild$ or schoolgirl$ or schoolboy$ or girl$ or boy$).tw.
31. Pediatrics/
32. p?ediatric$.tw.
33. or/27-32
34. 4 and 26 and 33
35. randomized controlled trial.pt.
36. controlled clinical trial.pt.
37. randomi#ed.ab.
38. placebo$.ab.

39. drug therapy.fs.
40. randomly.ab.
41. trial.ab.
42. groups.ab.
43. or/35-42
44. exp animals/ not humans.sh.
45. 43 not 44
46. 34 and 45

Lines 35 to 45 are the Cochrane highly sensitive search strategy for identifying randomized trials in MEDLINE (Ovid version)

3.13.2 Embase Search Strategy

Ovid platform; 1974-July 9, 2013

1. exp amphetamine/
2. (amphetamine$ or dexamphetamine$ or methamphetamine$ or dextroamphetamine$ or lisdexamphetamine$ or vyvanase$ or Dexedrin3 or desoxyn$ or adderall$).mp.
3. central stimulant agent/
4. 1 or 2 or 3
5. exp attention deficit disorder/
6. behavior disorder/
7. adhd.tw.
8. addh.tw.
9. adhs.tw.
10. adhs.tw.
11. "ad/hd".tw.
12. hyper-activ$.tw.
13. hyperactiv$.tw.
14. overactiv$.tw.
15. over-activ$.tw.
16. hyperkinesia/
17. hyperkin$.tw.
18. hyper-kin$.tw.
19. hkd.tw.
20. (minimal adj3 brain$ adj3 (damag$ or disorder$ or dysfunc$)).tw.
21. (attention$ adj3 (deficit$ or disorder$ or dysfunc$)).tw.
22. (behav$ adj3 (dysfunc$ or disorder$)).tw.
23. (behav$ adj3 (dysfunc$ or disorder$)).tw.
24. (impulsiv$ or inattentiv$ or inattention$).tw.
25. disruptiv$.tw.
26. or/5-25
27. child/

28. adolescent/
29. (adoles$ or teen$ or youth$ or young people or young person$).tw.
30. (child$ or toddler$ or preschool$ or pre-school or schoolchild$ or schoolgirl$ or schoolboy$ or girl$ or boy$).tw.
31. pediatrics/
32. p?ediatric$.tw.
33. or/27-32
34. 4 and 26 and 33
35. randomized controlled trial/
36. controlled clinical trial/
37. randomi#ed.ab.
38. placebo$.ab.
39. drug therapy/
40. randomly.ab.
41. trial.ab.
42. single blind procedure/
43. double blind procedure/
44. or/35-43
45. exp animal/ not human.sh.
46. 44 not 45
47. 34 and 46

3.13.3 PsycINFO Search Strategy
Ovid platform; 1806-July 9, 2013
1. exp amphetamine/
2. (amphetamine$ or dexamphetamine$ or methamphetamine$ or dextroamphetamine$ or lisdexamphetamine$ or vyvanase$ or Dexedrin3 or desoxyn$ or adderall$).mp.
3. exp attention deficit disorder/
4. behavior disorder/
5. adhd.tw.
6. addh.tw.
7. adhs.tw.
8. adhs.tw.
9. "ad/hd".tw.
10. hyper-activ$.tw.
11. hyperactiv$.tw.
12. overactiv$.tw.
13. over-activ$.tw.
14. hyperkin$.tw.
15. hyper-kin$.tw.
16. hkd.tw.

17. (minimal adj3 brain$ adj3 (damag$ or disorder$ or dysfunc$)).tw.
18. (attention$ adj3 (deficit$ or disorder$ or dysfunc$)).tw.
19. (behav$ adj3 (dysfunc$ or disorder$)).tw.
20. (behav$ adj3 (dysfunc$ or disorder$)).tw.
21. (impulsiv$ or inattentiv$ or inattention$).tw.
22. disruptiv$.tw.
23. (adoles$ or teen$ or youth$ or young people or young person$).tw.
24. (child$ or toddler$ or preschool$ or pre-school or schoolchild$ or schoolgirl$ or schoolboy$ or girl$ or boy$).tw.
25. pediatrics/
26. p?ediatric$.tw.
27. randomi#ed.ab.
28. placebo$.ab.
29. drug therapy/
30. randomly.ab.
31. trial.ab.
32. (doubl$ adj blind$).mp. [mp=title, abstract, heading word, table of contents, key concepts, original title, tests & measures]
33. (singl$ adj blind$).mp. [mp=title, abstract, heading word, table of contents, key concepts, original title, tests & measures]
34. 1 or 2
35. or/3-22
36. or/23-26
37. 34 and 35 and 36
38. or/27-33
39. exp animals/ not humans.sh.
40. 38 not 39
41. 37 and 40

3.13.4 CENTRAL Search Strategy

Ovid platform

1. exp Amphetamines/
2. (amphetamine$ or dexamphetamine$ or methamphetamine$ or dextroamphetamine$ or lisdexamphetamine$ or vyvanase$ or Dexedrin3 or desoxyn$ or adderall$).mp.
3. Central Nervous System Stimulants/
4. 1 or 2 or 3
5. exp Attention Deficit Disorder with Hyperactivity/
6. Child Behavior Disorders/
7. adhd.tw.
8. addh.tw.

9. adhs.tw.
10. adhs.tw.
11. "ad/hd".tw.
12. hyperactiv$.tw.
13. hyper-activ$.tw.
14. overactiv$.tw.
15. over-activ$.tw.
16. hyperkinesis/
17. hyperkin$.tw.
18. hyper-kin$.tw.
19. hkd.tw.
20. (minimal adj3 brain$ adj3 (damag$ or disorder$ or dysfunc$)).tw.
21. (attention$ adj3 (deficit$ or disorder$ or dysfunc$)).tw.
22. (behav$ adj3 (dysfunc$ or disorder$)).tw.
23. (behav$ adj3 (dysfunc$ or disorder$)).tw.
24. (impulsiv$ or inattentiv$ or inattention$).tw.
25. disruptiv$.tw.
26. or/5-25
27. exp child/
28. adolescent/
29. (adoles$ or teen$ or youth$ or young people or young person$).tw.
30. (child$ or toddler$ or preschool$ or pre-school or schoolchild$ or schoolgirl$ or schoolboy$ or girl$ or boy$).tw.
31. Pediatrics/
32. p?ediatric$.tw.
33. or/27-32
34. 4 and 26 and 33
35. randomized controlled trial.pt.
36. controlled clinical trial.pt.
37. randomi#ed.ab.
38. placebo$.ab.
39. randomly.ab.
40. trial.ab.
41. groups.ab.
42. or/35-41
43. exp animals/ not humans.sh. (3533521)
44. 42 not 43
45. 34 and 44

Chapter 4 Amphetamines and methylphenidate for pediatric ADHD: a systematic review and meta-analysis of N-of-1 evidence

Salima Punja, Dongying Xu, Christopher H Schmid, Lisa Hartling, Liana Urichuk, Catherine J Nikles, Sunita Vohra

Accepted by the *Journal of Clinical Epidemiology*

4.1 ABSTRACT

Objectives: To evaluate how data from N-of-1 trials may be used in systematic reviews and meta-analyses by examining the effects of amphetamine and methylphenidate for attention-deficit/hyperactivity disorder (ADHD).

Study Design and Setting: A systematic review and meta-analysis was conducted. An electronic search of MEDLINE, EMBASE, and PsychINFO for English language articles published from 1950-2013. N-of-1 trials of pediatric participants with a clinical diagnosis of ADHD that assessed either amphetamine or methylphenidate versus placebo were included. The primary outcome was improvement of core symptoms of ADHD. Studies with obtainable individual participant data were included in the meta-analysis. Weighted mean differences were computed using a random effects-model. Data were collected on total number of adverse events reported per participant.

Results: Nine studies were included in the amphetamine/placebo comparison, and ten in the methylphenidate/placebo comparison. Meta-analysis revealed a statistically significant difference in favor of amphetamine in 8/10 outcomes, and methylphenidate in 7/12 outcomes compared to placebo. A high degree of heterogeneity across participant treatment response was observed.

Conclusions: Although the focus of this study was assessing amphetamine and methylphenidate for pediatric ADHD, the central objective has general applications across treatment and disorder categories. By meta-analyzing N-of-1 evidence we were able to assess individual responses to treatment, as well as aggregate data across individuals and studies.

4.2 BACKGROUND

Systematic reviews have been proposed to be the highest form of evidence in healthcare (1). Systematic reviews attempt to identify and synthesize all high quality research evidence relevant to a particular clinical question, while remaining transparent and unbiased. Because randomized controlled trials (RCTs) are considered the gold standard in evidence-based medicine, the majority of systematic reviews are based on RCT evidence. Despite this, systematic reviews often overlook a specific subset of RCTs known as N-of-1 trials.

An N-of-1 trial is a multiple crossover trial performed in a single participant, often with randomization and blinding. N-of-1 trials provide an opportunity to determine the effect of an intervention on an individual who may not fit the eligibility criteria for an RCT. Although N-of-1 trials are primarily intended to evaluate therapeutic results in a single individual, preliminary data from systematic reviews indicate that the majority of published N-of-1 trials of health interventions comprise a series for the same condition-intervention pair (Chapter 2).

By virtue of their controlled methods (i.e. use of randomization, blinding and formal outcome assessment), N-of-1 trials may be worth systematically reviewing and meta-analyzing. In particular, when a series of N-of-1 trials have been conducted for the same condition-intervention pair, their meta-analysis may produce population treatment effects comparable to those yielded by RCTs. We aimed first to explore how data from N-of-1 trials may be used in systematic reviews and meta-analyses, in the context of a common

and important pediatric health condition, attention deficit/hyperactivity disorder (ADHD), in which numerous N-of-1 trials have been conducted to evaluate amphetamines and methylphenidate. Second, we aimed to examine the effect of amphetamines and methylphenidate on core symptoms of ADHD (inattention, impulsivity, hyperactivity) as measured by parent and teacher rating scales.

ADHD affects about 5% of children worldwide, making it among the most common pediatric neurodevelopmental disorders (2). ADHD is characterized by inattention, hyperactivity, and impulsivity, which negatively affect the child's socialization and education, and can have long-term ramifications throughout adulthood (3).

Psychostimulant drugs are recommended as the first-line of therapy for treating ADHD (3). Psychostimulants work by binding to the dopamine transporter in the brain, thereby blocking dopamine reuptake, and directly stimulate further dopamine release in the prefrontal cortex (4). It is thought that executive functioning is restored by increasing dopamine levels in the prefrontal cortex (4). Amphetamine and methylphenidate are the two most commonly prescribed stimulants for children with ADHD.

This review evaluates the effect of both amphetamines and methylphenidate for children with ADHD, an important and common pediatric condition. A considerable amount of published N-of-1 trial data for this condition has thus far never contributed to knowledge synthesis through systematic reviews and meta-analysis.

4.3. METHODS

4.3.1 Search strategy and selection of studies

We searched MEDLINE, EMBASE, and PsychINFO from 1950 to June, 2013. Relevant, published and unpublished N-of-1 trials were identified using the key terms: *attention deficit with hyperactivity disorder, child/adolescent/pediatric*, *amphetamine, methylphenidate, and single-patient experimental design.*

Only studies written in the English language were used. The reference lists of identified N-of-1 trials and review articles were screened to identify additional publications. Grey literature was identified by searching through the Network Digital Library of Theses and Dissertations, ProQuest Dissertations and Theses and Google Scholar. The complete search strategy for MEDLINE is provided in the Appendices 4.11.1-4.11.2 and was modified as appropriate for the other databases.

Selection of studies was based on screening of titles and/or abstracts independently by two authors (SP and DX). Both reviewers independently assessed the full-text articles of those studies whose inclusion was unclear based on abstracts alone. Final decisions were reached by consensus, with disagreements being resolved through discussion with the senior author (SV) as necessary.

4.3.2 Inclusion criteria

Prospectively planned N-of-1 trials (i.e. multiple crossover single-participant trials) were selected if they met the following criteria: i) participants were <18 years of age (we included studies that included both pediatric and adult participants if individual patient data (IPD) was available for the pediatric participants), with a clinical diagnosis of

ADHD as determined by DSM-III (5) or DSM-IV (3) criteria or equivalent; and ii) the trial compared an oral form of either amphetamine or methylphenidate to placebo. Only studies/authors that provided IPD were included in the meta-analysis. Only participants who had data on at least two treatment periods per intervention (i.e. at least two observations on amphetamine/methylphenidate and two observations on placebo) were included in the meta-analysis.

4.3.3 Data extraction

Two authors (SP and DX) independently extracted data from the included studies using predetermined data extraction forms. At the study level, extracted data included participant characteristics, interventions used, outcomes and trial design. At the individual participant level, extracted data included age, dose of intervention, gender, and individual responses to outcomes during each period. Discrepancies were resolved by a third party (SV).

4.3.4 Obtaining individual participant data

For those included studies that did not publish their IPD, authors were contacted up to three times to obtain this data.

4.3.5 Risk of bias assessment

For each included study, two reviewers (SP and DX) independently assessed risk of bias. The Cochrane Risk of bias tool was used to assess: random sequence generation, allocation concealment, blinding, incomplete outcome data, selective outcome reporting, and other sources of bias. Disagreements between the authors were resolved by a third party (SV).

4.3.6 Outcomes

The primary outcome of interest was efficacy defined as improvement of core symptoms of ADHD (inattention, impulsivity, hyperactivity) as measured by parent and/or teacher rating scales. Each core symptom as assessed by each rater (parent or teacher) was separately meta-analyzed. All data on adverse events were collected.

4.3.7 Analysis

The meta-analysis of continuous outcomes involved a two-step approach. Since most studies included >1 participant and therefore >1 N-of-1 trial, we calculated a mean and standard deviation (SD) for the intervention and a mean and SD for the placebo across the participant's measurements for each N-of-1 trial. A mean difference and SD of the mean difference was then calculated for each individual N-of-1 trial. The individual participant mean differences were then combined across all N-of-1 trials in a single study to get a weighted mean difference (WMD) using a random-effects model. We reported the total number of adverse events experienced per participant in his/her N-of-1 trial.

4.3.8 Subgroup analyses

We performed subgroup analyses according to sex (male versus female) and age (<13 years versus ≥13 years).

All calculations were performed using the Cochrane Collaboration's Review Manager Software (RevMan 2008).

4.4 RESULTS: AMPHETAMINES

4.4.1 Study selection and characteristics

The search of the electronic databases retrieved 57 studies. An additional six studies were retrieved from searching grey literature. After eliminating duplicates, 45 studies were identified for further consideration. After screening the titles and available abstracts, 17 studies were considered for possible inclusion. Of those 17 studies, 9 met the criteria for inclusion in the review (7-15). The flow of studies through the screening process of the review is shown in Figure 4.10.1 and is reported based on the PRISMA guidelines (6). The number of participants per study ranged between 1 and 54, with a total of 78 participants of whom 63 were boys and 11 were girls. The age of the participants ranged from 4-16 years. Characteristics of included studies are presented in Table 4.9.1.

Of the 9 included studies, only two studies were included in the meta-analysis. The Duggan et al study was a pilot study of the Nikles et al study; therefore they shared the same protocols. Two studies no longer had their IPD available, three other studies did not assess relevant outcomes for the meta-analysis; and an additional two studies had IPD available, however, each only involved one participant and assessed unique outcomes which could not be combined with any other studies. Useable IPD were obtained for 39 participants to be included in the meta-analysis.

4.4.2 Risk of Bias

Most studies failed to adequately describe methods used to generate and conceal the allocation sequence. Two studies were assessed as having a *high* risk of bias with respect to the method of sequence generation since both studies applied a predetermined sequence to all of their participants (9, 13). The majority of the studies failed to describe

blinding of participants, personnel and outcome assessors, resulting in a mostly *unclear* risk of bias assessment with respect to those domains. Most trials sufficiently addressed their incomplete outcome data, except one study that failed to report on reasons for withdrawal for all 20% of participants who did not complete their trials (8). Although protocols were not available for any of the studies, selective reporting was assessed alongside the methods. Most studies were rated as *unclear* in the 'other sources of bias' domain due to lack of clarity as to whether their outcome tools were valid. Details of the risk of bias assessment can be found in Table 4.9.2.

4.4.3 Meta-Analysis

Only two studies had useable IPD to be included in the meta-analyses of the primary outcome (7, 8). We obtained IPD on 37 participants from the Nikles et al study and two participants from the Duggan et al study. Both studies used parent and teacher versions of the DuPaul and Conners rating scales (Conners-Wells Adolescent Rating Scales for children >12 years) in order to assess core symptoms of ADHD. Forest plots are shown separately for each of the relevant outcomes (Figures 4.10.2.1 to 4.10.2.10).

a. Inattention (DuPaul rating scale) (a) Teacher reports: Data were obtained for 25 N-of-1 trials that used the teacher DuPaul rating scale to report on inattention. The meta-analysis revealed significant differences in favor of amphetamine (WMD of -4.35 (95%CI -6.30 to -2.41; $p<0.0001$)) (Figure 4.10.2.1). (b) Parent reports: Data were obtained for 30 N-of-1 trials that used the parent DuPaul rating scale to assess inattention. The meta-analysis showed significant results in favour of amphetamine with a WMD of -3.31 (95% CI -5.59 to -0.94; $p=0.006$) (Figure 4.10.2.2).

b. Inattention (Conners rating scale) (a) Teacher reports: Twenty-six N-of-1 trials

contributed to the meta-analysis of inattention on the Conner's rating scale. The meta-analysis revealed a statistically significant difference in favor of amphetamine with a WMD of -1.11 (95%CI -1.69 to -0.53; p=0.0002) (Figure 4.10.2.3). (b) Parent reports: Data were obtained for 27 N-of-1 trials that measured inattention using the parent Conners rating scale. The meta-analysis yielded no significant differences between amphetamine and placebo with a WMD of -2.59 (95%CI -5.35 to 0.18; p=0.07) (Figure 4.10.2.4).

c. Hyperactivity/Impulsivity (DuPaul scale) (a) Teacher reports: Data were obtained for 25 N-of-1 trials that used the DuPaul ADHD rating scale to assess hyperactivity/impulsivity. The meta-analysis revealed no significant differences between amphetamine and placebo with a WMD of -3.38 (95%CI -5.35 to -1.41; p=0.0008) (Figure 4.10.2.5). (b) Parent reports: Thirty N-of-1 trials contributed to this meta-analysis, which revealed a statistically significant WMD in favor of amphetamine at -3.23 (95%CI -5.00 to -1.46; p=0.0004) (Figure 4.10.2.6).

d. Hyperactivity (Conners rating scale) (a) Teacher reports: Twenty-three trials assessed hyperactivity/impulsivity using the Conners Teacher Rating Scale rating scale measured by teachers. The meta-analysis showed a statistically significant result in favor of amphetamine with a WMD of -2.42 (95%CI -3.68 to -1.16; p=0.0002) (Figure 4.10.2.7). (b) Parent reports: Data were obtained from 26 N-of-1 trials that used parent reports on the Conners rating scale to measure hyperactivity/impulsivity. The meta-analysis showed a statistically significant difference in favour of amphetamine with a WMD of -2.34 (95%CI -4.39 to -0.29; p=0.03) (Figure 4.10.2.8).

e. ADHD Index (Conners rating scale) (a) Teacher reports: Data were obtained on 25 N-

of-1 trials for this meta-analysis which showed a significant difference in favour of amphetamine (WMD -4.85 (95%CI -8.01 to -1.68; p=0.003) (Figure 4.10.2.9). (b) Parent reports: Twenty-six N-of-1 trials used parent ratings on the Conners scale to measure the ADHD index. The meta-analysis revealed a statistically significant difference in favor of placebo with a WMD of -5.67 (95%CI 10.33 to -1.01; p=0.02) (Figure 4.10.2.10).

4.4.4 Subgroup analysis

Subgroup analyses were performed based on sex (male versus female) and by age (child (<13 years) versus adolescent (≥13 years)).

Sex (male versus female)

We observed that males had statistically significant mean differences in favor of amphetamine on most outcomes, while females had no significant outcomes. Meaningful comparisons cannot be made given low amount of data on females (Table 4.9.3).

Age (<13 years versus ≥13 years)

We found children <13 years had statistically significant mean differences in favour of amphetamine on call outcomes, whereas children ≥13 years showed statistically significant differences in favour of amphetamine on only two outcomes. As above, meaningful comparisons cannot be made given the low amount of data on ≥13 year-olds (Table 4.9.4).

4.4.5 Adverse events

We obtained IPD on adverse events from 37 N-of-1 trials. Amphetamines presented with slightly more total reported adverse events (n=139) as compared with placebo (n=111) (see Table 4.9.5).

4.5 RESULTS: METHYLPHENIDATE

4.5.1 Study selection and characteristics

The search of the electronic databases retrieved 95 studies. An additional 3 studies were retrieved from searching grey literature. After eliminating duplicates, 62 studies were identified for further consideration. After screening titles and available abstracts, 30 full-text articles were assessed for eligibility. Of those 30 studies, 10 met the criteria for inclusion into the review (8, 11-13,15-20). The flow of studies through the screening process is shown in Figure 4.10.3. The number of participants per study ranged from 1 to 48, with a total of 71 participants of which 56 were boys and 15 were girls. The age of the participants ranged from 4-16 years. Characteristics of included studies are presented in Table 4.9.6.

Of the 10 included studies, only three studies were included in the meta-analysis. Four studies no longer had their IPD available, two studies did not assess outcomes relevant for the meta-analysis and one study author failed to respond to our data requests. Useable IPD were obtained for 39 participants to be included in the meta-analysis.

4.5.2 Risk of bias

All the studies were assessed as *unclear* under methods used to generate and conceal the allocation sequence, except for one study which had a high risk of bias for random sequence generation. The majority of the studies failed to describe blinding of participants, personnel and outcome assessors, resulting in a mostly *unclear* risk of bias assessment with respect to those domains. Most trials sufficiently addressed their incomplete outcome data; however, the study that contributed the largest amount of data

to the meta-analysis was rated *high* under this domain, since it did not discuss the reasons for withdrawal for all 20% of participants who dropped out. Although protocols were not available for any of the studies, selective reporting was assessed alongside methods, and most studies rated *low* in this domain. Most studies were rated as *unclear* under the 'other sources of bias' domain given their lack of clarity as to which ADHD diagnosis criteria were used as well as whether their outcome tools were valid. Details of the risk of bias assessment can be found in Table 4.9.7.

4.5.3 Meta-analysis

Only three studies had useable IPD to be included in the meta-analyses of the primary outcome (8, 15, 16). We obtained IPD on 36 participants from the Nikles et al study, two participants from the Zwaigenbaum et al study, and one participant from the Payton et al study. Nikles used parent and teacher versions of the DuPaul and Conners rating scales in order to assess the core symptoms of ADHD separately. Zwaigenbaum used parent and teacher versions of the Conners Abbreviated Rating scale to obtain total scores of impairment. Payton et al used teacher version of the Conner's behaviour rating scale to obtain total scores of impairment. Forest plots are shown separately for each of the relevant outcomes (Figure 4.10.4.1 to 4.10.4.12).

a. Inattention (DuPaul rating scale) (a) Teacher reports: Data were obtained for 30 N-of-1 trials that used the teacher DuPaul rating scale to report on inattention. The meta-analysis revealed significant differences in favour of methylphenidate (WMD of -4.67 (95%CI -6.79 to -2.56; $p<0.0001$)) (Figure 4.10.4.1). (b) Parent reports: Data were obtained for 30 N-of-1 trials that used the parent DuPaul rating scale to assess inattention. The meta-analysis showed significant results in favour of methylphenidate with a WMD

of -2.78 (95%CI -4.47 to -1.08; p<0.00001) (Figure 4.10.4.2).

b. Inattention (Conners rating scale) (a) Teacher reports: Thirteen N-of-1 trials contributed to the meta-analysis of inattention on the Conners rating scale. The meta-analysis revealed a statistically significant difference in favour of methylphenidate (WMD of -1.50 (95%CI -2.90 to -0.10; p=0.04) (Figure 4.10.4.3). (b) Parent reports: Data were obtained for 12 N-of-1 trials that measured inattention using the parent Conners rating scale. The meta-analysis showed no significant difference between methylphenidate and placebo (WMD of -0.73 (95%CI -1.82 to 0.36; p=0.19) (Figure 4.10.4.4).

c. Hyperactivity/Impulsivity (DuPaul scale) (a) Teacher reports: Data were obtained for 30 N-of-1 trials that used the DuPaul ADHD rating scale to assess hyperactivity/impulsivity. The meta-analysis revealed significant results in favor of methylphenidate with a WMD of -3.39 (95%CI -5.16 to -1.62; p<0.00001) (Figure 4.10.4.5). (b) Parent reports: Thirty N-of-1 trials contributed to this meta-analysis, which revealed a statistically significant WMD in favor of methylphenidate at -2.98 (95%CI -4.32 to -1.65; p<0.0001) (Figure 4.10.4.6).

d. Hyperactivity (Conners rating scale) (a) Teacher reports: Twelve trials were assessed hyperactivity/impulsivity using Conners rating scale measured by teachers. The meta-analysis showed a statistically significant result in favour of methylphenidate with a WMD of -1.22 (95%CI -2.11 to -0.33; p=0.007) (Figure 4.10.4.7). (b) Parent reports: Data were obtained from 12 N-of-1 trials that used parent reports on the Conner's rating scale to measure hyperactivity/impulsivity. The meta-analysis showed no significant differences between methylphenidate and placebo with a WMD of -0.65 (95%CI -2.24 to

0.94; p=0.42) (Figure 4.10.4.8).

e. ADHD Index (Conners rating scale) (a) Teacher reports: Data were obtained on 12 N-of-1 trials for this meta-analysis which showed a significant difference in favour of methylphenidate (WMD -3.00 (95%CI -5.15 to -0.84; p=0.007) (Figure 4.10.4.9). (b) Parent reports: Twelve N-of-1 trials used parent ratings on the Conner's scale to measure the ADHD index. The meta-analysis revealed no significant differences between methylphenidate and placebo with a WMD of -3.14 (95%CI -6.78 to -0.50; p=0.09) (Figure 4.10.4.10).

f. *Total scores* (a) Teacher reports: Data were meta-analyzed from 3 N-of-1 trials, which revealed no difference between methylphenidate and placebo (WMD -5.26 (95%CI -15.96 to 5.45; p=0.34)) (Figure 4.10.4.11). (b) Parent reports: Data were meta-analyzed from 2 N-of-1 trials, which revealed no significant difference between methylphenidate and placebo (WMD -0.45 (95%CI -2.38 to 1.48; p=0.65)) (Figure 4.10.4.12).

4.5.4 Subgroup analysis

Subgroup analyses were performed based on sex (male versus female) and by age (child (<13 years) versus adolescent (≥13 years)).

Sex (male versus female)

We observed that males had statistically significant mean differences in favour of amphetamine on all outcomes, while females had no significant outcomes. Meaningful comparisons cannot be made given the low amount of data on females (Table 4.9.8).

Age (<13 years versus ≥13 years)

We found children <13 years had statistically significant mean differences in favour of amphetamine on all outcomes, whereas children ≥13 years showed statistically

significant differences in favour of amphetamine on only one outcome. As above, meaningful comparisons cannot be made given the low amount of data on ≥13 year-olds (Table 4.9.9).

4.5.6 Adverse events

We obtained data on adverse events from 36 N-of-1 trials. Methylphenidate presented with slightly more total reported adverse events (n=128) as compared with placebo (n=117) (see Table 4.9.10).

4.6 DISCUSSION

In this study, we attempted to synthesize N-of-1 trials in a meta-analysis similar to meta-analysis of group RCTs. To our knowledge, a systematic review and meta-analysis of this kind has not previously been conducted. Although the focus of this study was assessing amphetamines and methylphenidate for pediatric ADHD, the central thesis may have general applications across treatment and disorder categories. By meta-analyzing only N-of-1 evidence we were able to assess individual responses to treatment, as well as aggregate data across individuals and studies. Our findings indicate that both amphetamines and methylphenidate were superior to placebo at the aggregate level on most outcomes including teacher ratings of inattention, and hyperactivity/impulsivity, as well as parent ratings of inattention and hyperactivity/impulsivity. These results are consistent with previously conducted meta-analyses (21), however, what the latter fail to depict is the variability of response to treatment at the individual participant level. It is important to note that our meta-analyses show a fair amount of heterogeneity across participants with some doing worse on stimulants, some doing better on stimulants, and some showing no difference between stimulant and placebo. These striking individual

differences reaffirm the importance of single-subject designs, such as the N-of-1 trial, which assesses treatment efficacy at the individual patient level, since aggregate data tend to mask individual variability in treatment response.

It has been noted that meta-analyses based on IPD provide the most comprehensive and reliable means of assessing the results of RCTs and are thus considered to be the 'gold standard' of systematic reviews (22). Benefits of IPD meta-analyses include the ability to explore more subgroups with the data, combine different scales of measurement, combat poor reporting and allow in-depth exploration of patient factors (23). As a result, the number of IPD meta-analyses has risen significantly from only 57 published articles before 2000 to an average of 49 articles published a year between 2005 and 2009 (24). Since N-of-1 trials can be viewed as IPD, meta-analysis of N-of-1 trials fits well with the movement towards IPD meta-analysis, where each participant contributes information regarding treatment effect. The ability to explore more subgroups with the data can be valuable, and allows us to estimate how patient characteristics modify treatment response. As the trend towards IPD meta-analyses continues to grow, single-subject research, particularly N-of-1 trials, can contribute a rich source of data allowing for more powerful and reliable assessments of treatment effects.

Single-subject research faces similar challenges as group-research. Just as in group-design systematic reviews and meta-analyses, N-of-1 trial research synthesis should be objective, replicable and valid. Flaws in their design, conduct, analysis, and reporting can cause the effect of an intervention to be underestimated or overestimated. This study was

limited by the quality of studies included. It was evident from the risk of bias assessment that the majority of trials were evaluated as being 'unclear' in the majority of the Risk of Bias domains. This suggests issues related to inadequate reporting of the N-of-1 studies. The use of reporting guidelines, such as the CONSORT statement (25), which was designed for reporting conventional RCTs, would be useful for N-of-1 trials. The CONSORT Extension for N-of-1 Trials (CENT) is a comprehensive checklist designed specifically for reporting of N-of-1 trials and is currently under review. Its adoption would be helpful to improve the quality of reporting of N-of-1 trials. Furthermore, since no risk of bias tool exists for N-of-1 trials, our study used the Cochrane Risk of Bias tool; however, this tool was designed primarily for parallel group RCTs. The development of a new risk of bias tool or modification of an existing tool is needed in order to accurately assess risk of bias in N-of-1 trials. Another limitation of our review is that it is likely not all N-of-1 trials are published or readily available. A recent review found only 108 trials (reporting on 2154 patients) over a 25-year time frame (26). One way of capturing these trials would be to encourage authors to register their N-of-1 protocols into an electronic repository (such as is done for conventional RCTs) in order to reduce selective outcome reporting and publication bias. Finally, one of the most significant limitations to our review was that IPD was not available for the majority of the studies, and therefore a large amount of data was excluded from our meta-analysis. In addition, data from one study contributed to most of our analysis, thereby potentially decreasing the generalizability of the results.

Meta-analysis has proven useful for investigating treatment effects. These efforts have

included mostly group-design studies; however, the volume of single-subject research has grown significantly, and should be included. Group-design outcome analyses have been criticized because they tend to mask individual variability to treatment effects and are not representative of clinical practice (27). N-of-1 trials provide individual response to treatment. Furthermore, when a series of N-of-1 trials of the same intervention is conducted in similar patients, with identical outcome measures, the results may be pooled for meta-analysis and provide a group mean effect.

4.7 ACKNOWLEDGEMENTS

The authors would like to thank Margaret Sampson for helping to put together the search strategies for this systematic review. We would also like to thank the authors who shared their individual patient data with us.

4.8 REFERENCES

1. Cook DJ, Mulrow CD, Haynes RB. Systematic Reviews: Synthesis of Best Evidence for Clinical Decisions. *Annals of Internal Medicine* 1997;126(5): 376-380. doi:10.7326/0003-4819-126-5-199703010-00006

2. Polanczyk G, Silva de Lima M, Horta BL, Biederman J, Rohde LA. Worldwide prevalence of ADHD. *American Journal of Psychiatry* 2007;4(164):942-948.

3. American Psychiatric Association. Diagnostic and Statistical Manual of Mental Disorders (DSM-IV-TR). 4th text revision edition. Washington, DC: American Psychiatric Association, 2000.

4. Arnsten AFT. Stimulants: therapeutic actions in ADHD. *Neuropsychopharmacology* 2006;31:2376-2383.

5. American Psychiatric Association. Diagnostic and Statistical Manual of Mental Disorders (DSM-III-R). 3rd edition.Washington, DC: American Psychiatric Association, 1987.

6. Moher D, Liberati A, Tetzlaff J, Altman DG, The PRISMA Group. *P*referred *R*eporting *I*tems for *S*ystematic Reviews and *M*eta-*A*nalyses: The PRISMA Statement. *PLoS Medicine* 2009;6(6):e1000097. doi:10.1371/journal.pmed1000097

7. Duggan CM, Mitchell G, Nikles CJ, Glasziou PP, Del Mar CB, Clavarino A. Managing ADHD in general practice: N-of-1 trials can help! *Australian Family Physician* 2000;29(12):1205-1209.

8. Nikles CJ, Mitchell GK, Del Mar CB, Clavarino A, McNairn N. An N-of-1 trial service in clinical practice: testing the effectiveness of stimulants for attention-deficit hyperactivity disorder. *Pediatrics* 2006;117(6):2040-2046.

9. Porrino LJ, Rapoport JL, Behar D, Ismond DR, Bunney Jr. WE. A naturalistic assessment of motor activity of hyperactive boys (Stimulant drug effects). *Archives of General Psychiatry* 1983;40:688-693.

10. LaRue Jr. RH, Northup J, Baumeister AA, Hawkins MF, Seale L, Williams T, Ridgway A. An evaluation of stimulant medication on the reinforcing effects of play. *Journal of Applied Behavior Analysis* 2008;41(1):143-147.

11. Speltz ML, Varley CK, Peterson K, Beilke RL. Effects of dextroamphetamine and contingency management on a preschooler with ADHD and oppositional defiant disorder. *American Academy of Child and Adolescent Psychiatry* 1988;27(2):175-178.

12. Hupp SD, Reitman D, Northup J, O'Callaghan P, LeBlanc M. The effects of delayed rewards, tokens, and stimulant medication on sportsmanlike behavior with ADHD-

diagnosed children. *Behavior Modification* 2002;26(2):148-162.

13. Neef NA, Bicard DF, Endo S, Coury DL, Aman MG. Evaluation of pharmacological treatment of impulsivity in children with attention deficit hyperactivity disorder. *Journal of Applied Behavior Analysis* 2005;38(2):135-146.

14. Kamien M. The use of an N-of-1 randomized clinical trial in resolving therapeutic doubt. *Australian Family Physician* 1998;27(Suppl 2):S103-S105.

15. Payton JB, Burkhart JE, Hersen M, Helsel WJ. Treatment of ADDH in mentally retarded children: A preliminary study. *American Academy of Child and Adolescent Psychiatry* 1989;28(5):761-767.

16. Zwaigenbaum L, Dick P, Handley-Derry M, Malone M, Jacobson S. "N-of-1" trials of methylphenidate in two children with Williams Syndrome and attention deficit hyperactivity disorder. *Journal of Developmental and Physical Disabilities* 2006;18(1):45-56.

17. Helsel WJ, Hersen M, Lubetsky MJ, Fultz SA, Sisson L, Harlovic CH. Stimulant drug treatment of four multihandicapped children using randomized single-case design. *Journal of the Multihandicapped Person* 1989;2(2):139-153.

18. Kutcher SP. Assessing and treating attention deficit disorder in adolescents. The clinical application of a single-case research design. *The British Journal of Psychiatry* 1986;149(6):710-715.

19. Reitman D, Hupp SDA, O'Callaghan PM, Gulley V, Northup J. The influence of a token economy and methylphenidate on attentive and disruptive behavior during sports with ADHD-diagnosed children. *Behavior Modification* 2001;25(2):305-323.

20. Anderson EE, Clement PW, Oettinger Jr. L. Methylphenidate compared with behavioral self-control in attention deficit disorder: Preliminary report. D*evelopmental and Behavioral Pediatrics* 1981;2(4):137-140.

21. Schacter HM, Pham B, King J, Langford S, Moher D. How efficacious and safe is short-acting methylphenidate for the treatment of attention-deficit disorder in children and adolescents? A meta-analysis. *Canadian Medical Association Journal* 2001;165(11):1475-1488.

22. Chalmers I. The Cochrane Collaboration: preparing, maintaining and disseminating systematic reviews of the effects of health care. *Annals of the New York Academy of Science* 1993;703:156-165.

23. Stewart LA, Tierney JF. To IPD or not to IPD? Advantages and disadvantages of systematic reviews using individual patient data. *Evaluation and Health Professions*

2002;25(1):76-97.

24. Riley RD, Lambert PC, Abo-Zaid G. Meta-analysis of individual participant data: Rationale, conduct, and reporting. *British Medical Journal* 2010;340:c221.

25. Schulz KF, Altman DG, Moher D, for the CONSORT Group. CONSORT 2010 Statement: updated guidelines for reporting parallel group randomised trials. *British Medical Journal* 2010;340:c332.

26. Gabler NB, Duan N, Vohra S, Kravtiz RL. N-of-1 trials in the medical literature: A systematic review. *Medical Care* 2011;49(8):761-767.

27. Rothwell PM. External validity of randomized controlled trials: "to whom do the results of this trial apply?" *Lancet*. 2005;365(9453):82-93.

4.9 TABLES

Table 4.9.1 Characteristics of included studies (amphetamine)

Study, year of publication	Country where study was conducted	Age (mean (SD)) in years	Number of participants (% male)	Dose of amphetamine	Study design	Length of treatment period (days)	Outcomes	IPD available
Porrino, 1983	United States	6-12 †	12 (100)	5-15 mg/day (adjusted individually)	ABAB	7	1. Continuous performance test 2. Core symptoms measured by Conners' Parent Questionnaire, and Abbreviated teacher rating scale)* 3. Adverse events* 4. Motor activity	No
Speltz, 1987	United States	4 (0)	1 (100)	5 mg/day and 10 mg/day‡	$AB_5B_{10}B_5B_{10}AB_5B_{10}$	5	1. Off-task behavior 2. Social behavior (onlooker, solitary, parallel, associative, and co-operative play) 3. Frequency of maladaptive behavior 4. Frequency of time-outs 5. Adverse events*	No
Payton, 1989	United States	7 (0)	1 (100)	5 mg/day and 10 mg/day‡	$AB_5B_{10}AB_5B_{10}$	5	1. Core symptoms (rated by Conners' behavior rating scale)* 2. Movement: any out-of-seat behavior 3. Ability to remain on-task	Yes (not utilized given unique outcome)

							for 10-second intervals	
Kamien, 1998	Australia	16 (0)	1 (100)	5 mg/day	ABAB	30	1. ADHD symptom score (unspecified tool)* 2. Patient's subjective experience	Yes (not utilized given unique outcome)
Duggan, 2000	Australia	11.5 (3.32)	4 (100)	Determined by physician and varied by participant	ABABAB	5	1. Core symptoms (inattention, hyperactivity, impulsivity)* 2. ADHD index* 3. Oppositional/Conduct problems 4. Adverse events*	Yes
Hupp, 2002	United States	6.5 (0.71)	2 (100)	10 mg/day	ABAB	NR	1. Sportsman-like behavior 2. Impact of delayed rewards 3. On-task behavior 4. Disruptive behavior	N/A
Neef, 2005	United States	10 (0)	1 (0)	20 mg/day	ABAB	7	1. Math problems: a. Percent attempted b. Percent correct c. Duration per session	N/A
Nikles, 2005	Australia	10.2 (2.25)	54 (80)	Determined by physician (ranged from 5-20.5 mg/day)	ABABAB	For first 32 participants: 5 For last 22 participants: 2	1. Core symptoms (inattention, hyperactivity, impulsivity)* 2. ADHD index* 3. Oppositional/Conduct problems 4. Adverse events*	Yes

LaRue, 2008	United States	4-6†	2 (100)	20 mg/day	AB (x10)	1	1. Play and social behavior measured by a reinforce assessment procedure	No

Abbreviations: IPD, individual patient data; SD, standard deviation; NR, not reported; N/A, not applicable

*Indicates outcomes relevant for meta-analysis

†Only age range reported (individual patient data not available)

‡Study assessed two doses of amphetamine

Table 4.9.2: Risk of bias assessment (amphetamine)

Study, year of publication	Random sequence generation	Allocation concealment	Blinding of participants and personnel	Blinding of outcome assessment	Incomplete outcome data	Selective reporting	Other sources of bias
Porrino, 1983	HIGH*	UNCLEAR	UNCLEAR	UNCLEAR	LOW	LOW	LOW
Speltz, 1987	UNCLEAR	UNCLEAR	UNCLEAR	UNCLEAR	LOW	LOW	UNCLEAR†
Payton, 1989	UNCLEAR	UNCLEAR	LOW	LOW	LOW	LOW	UNCLEAR [Report does not state which diagnostic criteria was used for participant inclusion]
Kamien, 1998	LOW	LOW	LOW	LOW	LOW	LOW	UNCLEAR†
Duggan, 2000	UNCLEAR	UNCLEAR	LOW	UNCLEAR	LOW	HIGH [data on behavioral outcomes mentioned in the methods are not reported on]	LOW
Hupp, 2002	UNCLEAR	UNCLEAR	UNCLEAR	UNCLEAR	LOW	LOW	UNCLEAR†
Neef, 2005	HIGH*	UNCLEAR	LOW	UNCLEAR	LOW	LOW	UNCLEAR†

Nikles, 2005	UNCLEAR	UNCLEAR	LOW	LOW	HIGH [Only 80% of N-of-1 trials commenced are reported on, authors do not report what happened to the other 20%]	HIGH [Data on behavioral outcomes mentioned in the methods are not reported on]	LOW
LaRue, 2008	UNCLEAR	UNCLEAR	UNCLEAR	UNCLEAR	LOW	LOW	UNCLEAR[†]

UNCLEAR: not adequately described

LOW: study adequately reported on domain and appears to be free of bias on the respective domain

[*] Sequence was not randomly generated

[†]Validity of outcome measurement tool used is unclear

Table 4.9.3: Subgroup analysis by sex (amphetamine)

Subgroup	Amphetamine		Placebo		Mean Difference
	No. of trials	N*	No. of trials	N*	Random, 95% CI
Outcome: Inattention by teacher ratings on DuPaul scale					
Male	21	58	21	58	-4.89 [-7.10, -2.67]†
Female	5	14	5	13	-1.31 [-4.20, 1.57]
Outcome: Inattention by parent ratings on DuPaul scale					
Male	24	70	24	70	-3.80 [-6.98, -0.62]†
Female	6	18	6	18	-0.21 [-1.32, 0.91]
Outcome: Hyperactivity-Impulsivity by teacher ratings on DuPaul scale					
Male	22	60	22	61	-3.74 [-5.93, -1.55]†
Female	4	11	4	10	0.57 [-0.65, 1.80]
Outcome: Hyperactivity-Impulsivity by parent ratings on DuPaul scale					
Male	24	70	24	70	-3.84 [-5.92, -1.76]†
Female	6	18	6	18	-0.85 [-3.55, 1.85]
Outcome: Inattention by teacher ratings on Conners scale					
Male	20	54	20	54	-1.13 [-1.72, -0.54]†
Female	6	17	6	18	-1.25 [-3.14, 0.63]
Outcome: Inattention by parent ratings on Conners scale					
Male	23	68	23	66	-2.78 [-6.15, 0.60]
Female	4	12	4	12	-0.94 [-2.16, 0.28]
Outcome: Hyperactivity by teacher ratings on Conners scale					
Male	19	51	19	51	-2.64 [-4.09, -1.19]†
Female	4	11	4	12	-1.66 [-4.08, 0.77]
Outcome: Hyperactivity by parent ratings on Conners scale					
Male	22	65	22	63	-2.58 [-5.02, -0.14]†
Female	4	12	4	12	-0.44 [-1.22, 0.34]
Outcome: ADHD Index by teacher ratings on Conners scale					
Male	20	55	20	55	-4.97 [-8.81, -1.12]†
Female	5	14	5	15	-3.80 [-7.68, 0.09]
Outcome: ADHD Index by parent ratings on Conners scale					
Male	23	68	23	66	-6.15 [-11.32, -0.99]†
Female	3	9	3	9	-1.58 [-6.26, 3.09]

*refers to total number of observations per trial per intervention arm which in this review are used to reflect number of participants per intervention arm

†significant results in favor of amphetamine

Table 4.9.4: Subgroup analysis by age (amphetamine)

Subgroup	Amphetamine		Placebo		Mean Difference
	No. of trials	N*	No. of trials	N*	Random, 95% CI
Outcome: Inattention by teacher ratings on DuPaul scale					
<13 years	22	61	21	60	-5.44 [-7.85, -3.03]†
≥13 years	4	11	4	11	0.73 [-1.32, 2.79]
Outcome: Inattention by parent ratings on DuPaul scale					
<13 years	27	79	27	80	-3.77 [-6.35, -1.20]†
≥13 years	3	9	3	8	0.52 [-5.17, 6.21]
Outcome: Hyperactivity-Impulsivity by teacher ratings on DuPaul scale					
<13 years	22	60	21	60	-4.08 [-6.70, -1.46]†
≥13 years	4	11	4	11	-0.64 [-2.33, 1.04]
Outcome: Hyperactivity-Impulsivity by parent ratings on DuPaul scale					
<13 years	27	79	27	80	-3.26 [-5.16, -1.37]†
≥13 years	3	9	3	8	-1.41 [-4.38, 1.57]
Outcome: Inattention by teacher ratings on Conners scale					
<13 years	20	55	20	56	-1.06 [-1.77, -0.35]†
≥13 years	6	16	6	16	-1.25 [-2.32, -0.18]†
Outcome: Inattention by parent ratings on Conners scale					
<13 years	22	65	22	64	-2.73 [-5.00, -0.45]†
≥13 years	5	15	5	14	0.07 [-2.37, 2.51]
Outcome: Hyperactivity by teacher ratings on Conners scale					
<13 years	17	46	17	47	-3.11 [-4.86, -1.36]†
≥13 years	6	16	6	16	-0.78 [-1.56, 0.00]
Outcome: Hyperactivity by parent ratings on Conners scale					
<13 years	21	62	21	61	-2.73 [-5.00, -0.45]†
≥13 years	5	15	5	14	0.07 [-2.37, 2.51]
Outcome: ADHD Index by teacher ratings on Conners scale					
<13 years	19	53	19	54	-5.23 [-9.26, -1.21]†
≥13 years	6	16	6	16	-2.84 [-5.60, -0.07]†
Outcome: ADHD Index by parent ratings on Conners scale					
<13 years	21	62	21	61	-5.44 [-9.31, -1.56]†
≥13 years	5	15	5	14	-6.92 [-21.05, 7.21]

*refers to total number of observations per trial per intervention arm which in this review are used to reflect number of participants per intervention arm

†significant results in favor of amphetamine

Table 4.9.5: Number of adverse events reported by individual participants (amphetamine)

Patient ID	Amphetamine	Placebo
JN4045	0	0
JN4057	3	3
JN4063	7	3
JN4053	7	7
JN4065	0	1
JN4064	9	12
JN4067	11	2
JN4069	4	3
JN4077	3	7
JN4082	12	10
JN5013	0	0
JN5017	1	0
JN5020	3	3
JN5023	12	12
JN4000	1	2
JN4005	6	5
JN4006	6	0
JN4010	2	1
JN4011	2	1
JN4012	1	1
JN4013	3	3
JN4014	1	1
JN4015	3	3
JN4020	7	2
JN4021	5	5
JN4024	2	2
JN4025	4	4
JN4029	0	0
JN4035	0	0
JN4036	0	0
JN4037	0	2
JN4039	10	10
JN4041	2	2
JN4042	7	2
JN4044	2	1
JN4048	3	1
Speltz1	0	15
Total	139	111

Table 4.9.6: Characteristics of included studies (methylphenidate)

Study, year of publication	Country where study was conducted	Age (mean (SD))	Number of N-of-1 trials started (% male)	Dose of methylphenidate (mg/day)	Study design	Length of treatment period (days)	Outcomes	IPD available on relevant outcomes
Anderson, 1981	United States	6-9†	4 (100)	Weight-based dosing (0.3 mg/kg)	ABAB	14	1. CTRS* 2. CPRS* 3. Academic performance (WISC-R), WRAT, Bender Motor Gestalt Test 4. Sustained attention, measured by CCT	No
Kutcher, 1986	United Kingdom	14	1 (100)	30	ABAB	4	1. Conners scale rated by nursing staff, physicians, teacher*	No
Helsl, 1989	United States	6.12 (1.25)	4 (50)	Weight-based dosing -3 doses assessed/participant: 0.3 mg/kg, 0.45 mg/kg, 0.6 mg/kg	$AB_1AB_2B_3B_3B_1$	4	1. On-task behavior 2. In-seat behavior 3. Talking-out behavior 4. Task accuracy 5. Duration of task completion 6. Abbreviated Conners Teacher Rating Scale* 7. Iowa Conners Teacher Rating Scale* 8. Affect related behaviors measured by Affect Assessment Scale	No
Payton, 1989	United States	7.5 (0)	1 (100)	20 and 30	ABCABC	5	1. Core symptoms (rated by Conners' behavior	Yes

							rating scale)* 2. Movement: any out-of-seat behavior 3. Ability to remain on-task for 10-second intervals	
Reitman, 2001	United States	6.3 (0.47)	3 (67)	Varied between participants (ranged from 10-15)	ABAB	NR	1. Attentive behavior 2. Disruptive behavior 3. ADHD index of CTRS*	No
Hupp, 2002	United States	6 (0)	1 (100)	7.5	ABAB	NR	1. Sportsman-like behavior 2. Impact of delayed rewards 3. On-task behavior 4. Disruptive behavior	N/A
Neef, 2005	United States	10.7 (2.5)	3 (67)	Determined by physician (ranged from 7.5-40)	ABAB	7	1. Math problems: a. Percent attempted b. Percent correct c. Duration per session	N/A
Nikles, 2005	Australia	11.03 (2.6)	48 (79)	Determined by physician (ranged from 7.5-40)	AB (x3)	For first 17 participants: 5 For last 31 participants: 2	1. Core symptoms (inattention, hyperactivity, impulsivity)* 2. ADHD index* 3.Oppositional/Conduct problems 4. Adverse events*	Yes
Zwaigenbaum, 2006	Canada	8 (1.41)[†]	3 (67)	7.5 mg/day	AB (x5)	7	1. Conners Abbreviated Rating Scale for Parents* 2. Conners Abbreviated	Yes

							Rating Scale for Teachers* 3. Adverse events*	
LaRue, 2008	United States	4-6‡	3 (100)	Varied between participants (ranged from 36-54)	AB (x10)	1	1. Play and social behavior measured by a reinforce assessment procedure	No

Abbreviations: IPD, individual patient data; SD, standard deviation; NR, not reported; N/A, not applicable; CTRS, Conners Teacher Rating Scale; CPRS, Conners Parent Rating Scale; WISC-R, Wechsler Intelligence Scale for Children-Revised; WRAT, Justak Wide Range Achievement Test; CCT, Children's Checking Test

*Indicates outcomes relevant for meta-analysis

†Age reported for 2/3 participants

‡Only age range reported (individual patient data not available)

Table 4.9.7: Risk of bias assessment (methylphenidate)

Study, year of publication	Random sequence generation	Allocation concealment	Blinding of participants and personnel	Blinding of outcome assessment	Incomplete outcome data	Selective reporting	Other sources of bias
Anderson, 1981	UNCLEAR	UNCLEAR	UNCLEAR	UNCLEAR	UNCLEAR	HIGH [Data on majority of outcomes reported in methods are not reported on in the results]	UNCLEAR*
Kutcher, 1986	UNCLEAR	UNCLEAR	UNCLEAR	UNCLEAR	LOW	LOW	UNCLEAR*
Helsl, 1989	UNCLEAR	UNCLEAR	LOW	UNCLEAR	UNCLEAR	LOW	UNCLEAR†
Payton, 1989	UNCLEAR	UNCLEAR	UNCLEAR	LOW	LOW	LOW	UNCLEAR*†
Reitman, 2001	UNCLEAR	UNCLEAR	UNCLEAR	LOW	LOW	LOW	LOW
Hupp, 2002	UNCLEAR	UNCLEAR	UNCLEAR	UNCLEAR	LOW	LOW	UNCLEAR†
Neef, 2005	HIGH [Sequence was not randomly generated]	UNCLEAR	LOW	UNCLEAR	LOW	LOW	UNCLEAR†
Nikles, 2005	UNCLEAR	UNCLEAR	LOW	LOW	HIGH [Only 80% of N-of-1 trials commenced are reported on,	HIGH [Data on behavioral outcomes mentioned in the methods are not	LOW

					authors do not discuss what happened to the other 20%]	reported on in the results]	
Zwaigenbaum, 2006	LOW	UNCLEAR	UNCLEAR	UNCLEAR	HIGH [no data is provided for one out of the three participants in this series]	LOW	LOW
LaRue, 2008	UNCLEAR	UNCLEAR	UNCLEAR	UNCLEAR	LOW	LOW	UNCLEAR[†]

UNCLEAR: domain was not adequately reported on

LOW: domain was adequately reported on and therefore study appears to be free of bias on the respective domain

[*]Report does not state which diagnostic criteria was used for ADHD diagnosis

[†]Validity of outcome measurement tool used is unclear

Table 4.9.8: Subgroup analysis by sex (methylphenidate)

Subgroup	Methylphenidate		Placebo		Mean Difference
	No. of trials	N*	No. of trials	N*	Random, 95% CI
Outcome: Inattention by teacher ratings on DuPaul scale					
Male	25	68	25	69	-4.81 [-7.18, -2.44]†
Female	5	15	5	14	-3.99 [-9.15, 1.16]
Outcome: Inattention by parent ratings on DuPaul scale					
Male	24	68	24	70	-3.09 [-5.10, -1.09]†
Female	6	16	6	18	-1.46 [-4.00, 1.07]
Outcome: Hyperactivity-Impulsivity by teacher ratings on DuPaul scale					
Male	25	67	25	69	-3.78 [-5.90, -1.65]†
Female	5	15	5	14	-1.52 [-3.89, 0.84]
Outcome: Hyperactivity-Impulsivity by parent ratings on DuPaul scale					
Male	24	68	24	70	-3.48 [-5.10, -1.86]†
Female	6	17	6	18	-1.14 [-2.36, 0.09]

*refers to total number of observations per trial per intervention arm which in this review are used to reflect number of participants per intervention arm

†significant results in favor of methylphenidate

Table 4.9.9: Subgroup analysis by age (methylphenidate)

Subgroup	Methylphenidate		Placebo		Mean Difference
	No. of trials	N*	No. of trials	N*	Random, 95% CI
Outcome: Inattention by teacher ratings on DuPaul scale					
<13 years	23	63	23	64	-4.65 [-7.07, -2.22]†
≥13 years	7	20	7	19	-4.78 [-10.28, 0.72]
Outcome: Inattention by parent ratings on DuPaul scale					
<13 years	22	60	22	64	-2.90 [-4.87, -0.94]†
≥13 years	8	24	8	24	-2.42 [-6.43, 1.59]
Outcome: Hyperactivity-Impulsivity by teacher ratings on DuPaul scale					
<13 years	23	62	23	64	-3.31 [-5.11, -1.52]†
≥13 years	7	20	7	19	-3.25 [-9.52, 3.02]
Outcome: Hyperactivity-Impulsivity by parent ratings on DuPaul scale					
<13 years	22	61	22	64	-3.37 [-5.11, -1.64]†
≥13 years	8	24	8	24	-2.02 [-4.04, -0.01]†

*refers to total number of observations per trial per intervention arm which in this review are used to reflect number of participants per intervention arm

†significant results in favor of methylphenidate

Table 4.9.10: Number of adverse events reported by individual participants (methylphenidate)

Patient ID	Methylphenidate	Placebo
JN4049	5	0
JN4054	3	3
JN4060	7	6
JN4061a	1	1
JN4066	0	0
JN4072	0	0
JN4073	2	5
JN4078	3	3
JN4080	0	0
JN5001	0	0
JN5002	10	8
JN5003	4	3
JN5004	8	7
JN5005	2	4
JN5006	1	6
JN5010	4	4
JN5011	5	5
JN5012	7	6
JN5014	9	11
JN5015	10	11
JN5018	3	2
JN5026	4	6
JN5027	0	0
JN4001	9	1
JN4004	7	2
JN4008	4	4
JN4018	1	0
JN4019	3	3
JN4023	0	0
JN4026	0	0
JN4027	3	3
JN4032	0	0
JN4033	4	6
JN4034	7	0
JN4043	2	7
JN4046	0	0
Total	128	117

4.10 FIGURES

Figure 4.10.1 Flow of studies (amphetamine)

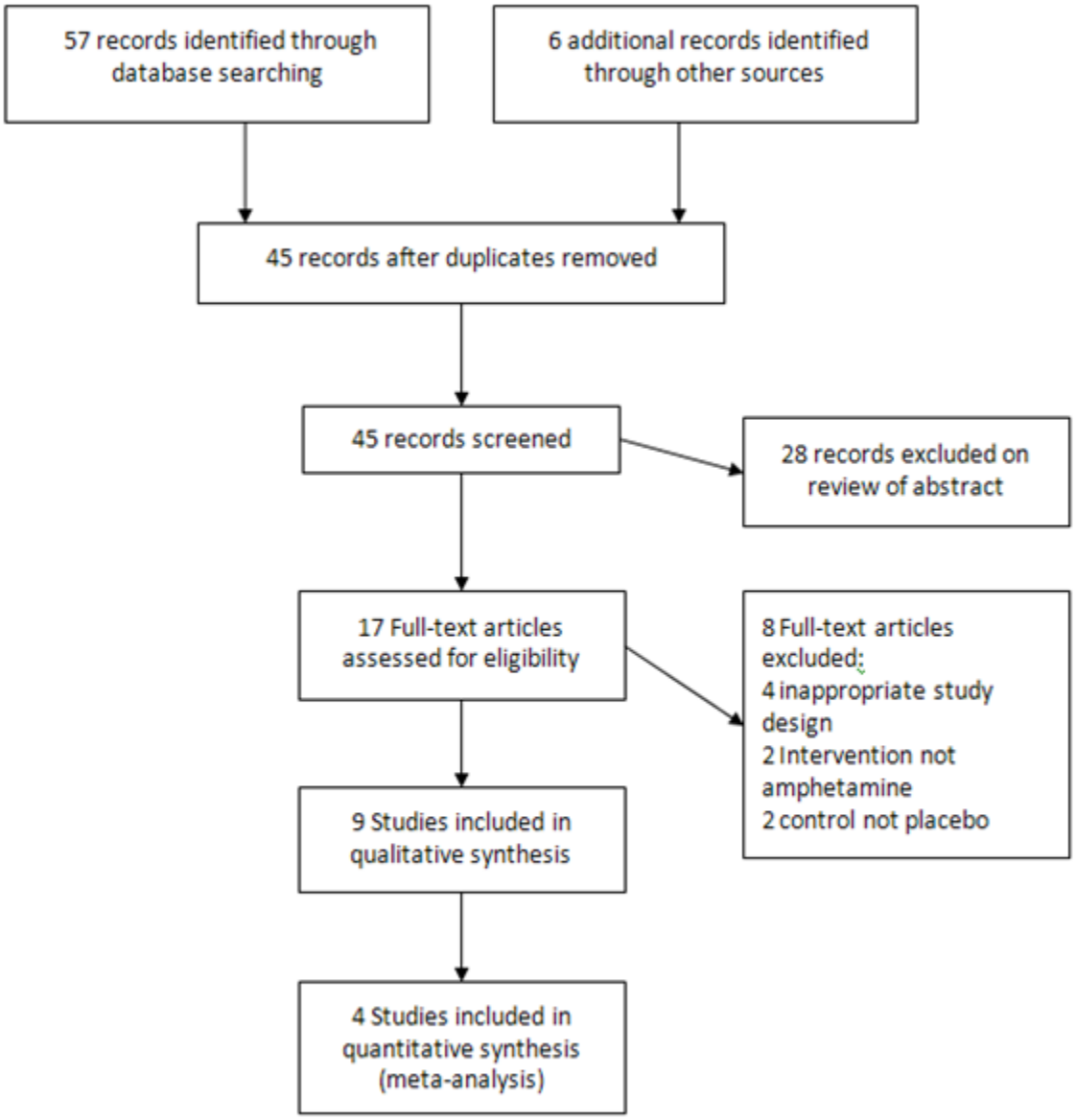


Figure 4.10.2: Meta-analysis (amphetamine)

Figure 4.10.2.1 Teacher ratings of inattention on the DuPaul scale

Study or Subgroup	Amphetamine Mean	SD	Total	Placebo Mean	SD	Total	Weight	Mean Difference IV, Random, 95% CI
JN4000	4	3	2	14	2.88	2	3.4%	-10.00 [-15.76, -4.24]
JN4006	8.33	4.71	3	13.67	1.89	3	3.4%	-5.34 [-11.08, 0.40]
JN4011	16	2	2	12.5	0.5	2	4.4%	3.50 [0.64, 6.36]
JN4012	27.67	0.47	3	28.5	0.5	2	4.8%	-0.83 [-1.70, 0.04]
JN4014	18.5	1.5	3	16.5	1.5	2	4.4%	2.00 [-0.68, 4.68]
JN4015	0.5	0.5	2	26.5	2.5	2	4.2%	-26.00 [-29.53, -22.47]
JN4020	18	2.23	3	19.67	1.37	3	4.4%	-1.67 [-4.63, 1.29]
JN4024	3	0.82	3	9	3.74	3	3.9%	-6.00 [-10.33, -1.67]
JN4025	29.5	0.5	2	27	3	2	4.0%	2.50 [-1.72, 6.72]
JN4029	8	4.32	3	14.33	3.77	3	3.1%	-6.33 [-12.82, 0.16]
JN4036	1	1.41	3	6.33	5.31	3	3.2%	-5.33 [-11.55, 0.89]
JN4037	0.67	1.15	3	0.33	0.58	3	4.7%	0.34 [-1.12, 1.80]
JN4038	1.33	2.31	3	3.67	3.51	3	3.8%	-2.34 [-7.09, 2.41]
JN4041	8	9.2	3	17	2.83	3	1.9%	-9.00 [-19.89, 1.89]
JN4042	4.67	0.47	3	15	5.1	3	3.4%	-10.33 [-16.13, -4.53]
JN4045	0.67	0.58	3	0.67	0.58	3	4.8%	0.00 [-0.93, 0.93]
JN4063	7.67	2.87	3	6.67	2.08	3	4.0%	1.00 [-3.01, 5.01]
JN4065	3.67	1.53	3	4	1	3	4.6%	-0.33 [-2.40, 1.74]
JN4067	11	4.08	3	20.67	2.06	3	3.6%	-9.67 [-14.84, -4.50]
JN4069	5.33	2.64	3	10	3.19	3	3.8%	-4.67 [-9.36, 0.02]
JN4077	9.5	1.5	2	15	2.94	3	4.1%	-5.50 [-9.42, -1.58]
JN4082	24.33	1.25	3	23.67	1.7	3	4.5%	0.66 [-1.73, 3.05]
JN5013	8	2.45	3	17.33	4.64	3	3.3%	-9.33 [-15.27, -3.39]
JN5017	6.67	2.69	3	17.33	1.23	3	4.2%	-10.66 [-14.01, -7.31]
JN5020	9.33	8.34	3	11.5	7.5	2	1.4%	-2.17 [-16.21, 11.87]
JN5023	2.5	0.5	2	8.67	1.7	3	4.6%	-6.17 [-8.21, -4.13]
Total (95% CI)			**72**			**71**	**100.0%**	**-4.35 [-6.30, -2.41]**

Heterogeneity: Tau² = 20.21; Chi² = 330.09, df = 25 (P < 0.00001); I² = 92%
Test for overall effect: Z = 4.40 (P < 0.0001)

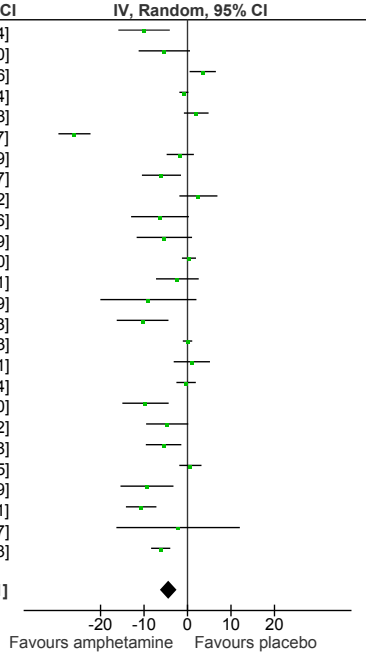


Figure 4.10.2.2 Parent ratings of inattention on the DuPaul scale

Study or Subgroup	Amphetamine Mean	SD	Total	Placebo Mean	SD	Total	Weight	Mean Difference IV, Random, 95% CI
JN4000	14	4.08	3	12.67	0.94	3	3.4%	1.33 [-3.41, 6.07]
JN4005	16.67	0.94	3	16	1.41	3	3.9%	0.67 [-1.25, 2.59]
JN4006	6	2.94	3	5.67	3.86	3	3.3%	0.33 [-5.16, 5.82]
JN4010	8	7	2	12	1	2	2.4%	-4.00 [-13.80, 5.80]
JN4011	2.33	1.7	3	3.33	2.36	3	3.7%	-1.00 [-4.29, 2.29]
JN4012	6	0.82	3	4	0.82	3	3.9%	2.00 [0.69, 3.31]
JN4013	5	1.41	3	13.67	3.3	3	3.5%	-8.67 [-12.73, -4.61]
JN4014	10	4.97	3	8	2.16	3	3.1%	2.00 [-4.13, 8.13]
JN4015	3	2.16	3	27	2.94	3	3.5%	-24.00 [-28.13, -19.87]
JN4020	7	1.63	3	8	2.94	3	3.6%	-1.00 [-4.80, 2.80]
JN4021	2.33	1.25	3	16.67	1.89	3	3.8%	-14.34 [-16.90, -11.78]
JN4024	9	5.72	3	14.33	2.49	3	2.9%	-5.33 [-12.39, 1.73]
JN4025	12.67	5.19	3	19.67	4.19	3	2.8%	-7.00 [-14.55, 0.55]
JN4029	12.67	3.09	3	11.67	1.25	3	3.6%	1.00 [-2.77, 4.77]
JN4035	0.67	0.47	3	0.67	0.58	3	3.9%	0.00 [-0.84, 0.84]
JN4036	3.33	2.06	3	7.67	5.91	3	2.9%	-4.34 [-11.42, 2.74]
JN4037	15	5.29	3	27	1.73	3	3.1%	-12.00 [-18.30, -5.70]
JN4039	16	4.97	3	16.33	5.91	3	2.6%	-0.33 [-9.07, 8.41]
JN4041	3.33	0.94	3	11.33	9.39	3	2.2%	-8.00 [-18.68, 2.68]
JN4042	9.33	4.19	3	11.33	3.09	3	3.2%	-2.00 [-7.89, 3.89]
JN4045	5.67	0.58	3	9.33	3.79	3	3.5%	-3.66 [-8.00, 0.68]
JN4063	12.67	1.53	3	7	1.41	2	3.8%	5.67 [3.06, 8.28]
JN4065	2	1.73	3	2.33	3.21	3	3.5%	-0.33 [-4.46, 3.80]
JN4067	3	1.63	3	20.67	1.7	3	3.8%	-17.67 [-20.34, -15.00]
JN4069	11.33	2.36	3	12	2.94	3	3.5%	-0.67 [-4.94, 3.60]
JN4077	17.5	0.5	2	16	5.35	3	3.1%	1.50 [-4.59, 7.59]
JN4082	16.67	2.06	3	17.67	3.3	3	3.5%	-1.00 [-5.40, 3.40]
JN5013	10	2.83	3	11	1.41	3	3.6%	-1.00 [-4.58, 2.58]
JN5020	21	2.16	3	18.33	6.13	3	2.9%	2.67 [-4.68, 10.02]
JN5023	18	0.82	3	19.33	3.3	3	3.6%	-1.33 [-5.18, 2.52]
Total (95% CI)			**88**			**88**	**100.0%**	**-3.31 [-5.69, -0.94]**

Heterogeneity: Tau² = 37.22; Chi² = 461.99, df = 29 (P < 0.00001); I² = 94%
Test for overall effect: Z = 2.73 (P = 0.006)

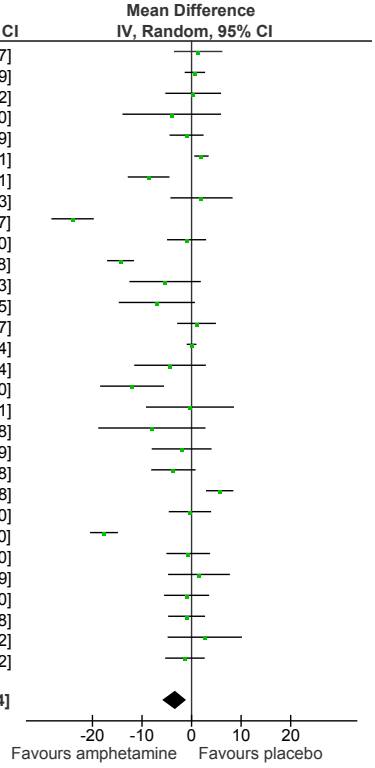


Figure 4.10.2.3 Teacher ratings of inattention on the Conners scale

Study or Subgroup	Amphetamine Mean	SD	Total	Placebo Mean	SD	Total	Weight	Mean Difference IV, Random, 95% CI
Duggan2	1.67	2.89	3	2	1	3	2.1%	-0.33 [-3.79, 3.13]
JN4000	3.5	0.5	2	5.5	1.5	2	3.9%	-2.00 [-4.19, 0.19]
JN4005	5	1	2	8	5.1	3	0.9%	-3.00 [-8.94, 2.94]
JN4006	6	2.16	3	7.67	1.25	3	2.8%	-1.67 [-4.49, 1.15]
JN4011	9.5	0.5	2	10.5	0.5	2	6.9%	-1.00 [-1.98, -0.02]
JN4012	7.67	0.47	3	8.5	0.5	2	7.2%	-0.83 [-1.70, 0.04]
JN4014	11.5	1.5	2	12	0.82	3	3.7%	-0.50 [-2.78, 1.78]
JN4015	0.5	0.5	2	7	2	2	2.8%	-6.50 [-9.36, -3.64]
JN4020	10.33	1.7	3	9.33	2.06	3	2.6%	1.00 [-2.02, 4.02]
JN4024	1.33	0.47	3	1.67	1.7	3	4.3%	-0.34 [-2.34, 1.66]
JN4025	10.5	1.5	2	8.67	0.47	3	4.0%	1.83 [-0.32, 3.98]
JN4029	5.67	0.47	3	7	0.82	3	6.7%	-1.33 [-2.40, -0.26]
JN4035	1.67	0.47	3	1.67	0.47	3	7.5%	0.00 [-0.75, 0.75]
JN4036	1	0.87	3	3.5	2.5	3	2.6%	-2.50 [-5.50, 0.50]
JN4037	0.67	0.58	3	1	1.32	3	5.1%	-0.33 [-1.96, 1.30]
JN4038	2.33	1.15	3	3	1	3	4.9%	-0.67 [-2.39, 1.05]
JN4039	2.17	2.02	3	1.5	0.71	2	3.3%	0.67 [-1.82, 3.16]
JN4042	2.67	0.47	3	4.5	0.5	3	7.5%	-1.83 [-2.61, -1.05]
JN4044	5.17	2.75	3	6.67	1.26	3	2.2%	-1.50 [-4.92, 1.92]
JN4045	2	1	3	2	1	3	5.2%	0.00 [-1.60, 1.60]
JN4048	10.5	0.5	2	12	1.41	2	4.1%	-1.50 [-3.57, 0.57]
JN4053	3.67	0.47	3	9	2.65	3	2.6%	-5.33 [-8.38, -2.28]
JN4063	4.33	3.11	3	8.33	2.37	3	1.4%	-4.00 [-8.42, 0.42]
JN4064	5.67	5.51	3	7.33	2.52	3	0.7%	-1.66 [-8.52, 5.20]
JN4067	5.33	4.99	3	14.67	0.47	3	0.9%	-9.34 [-15.01, -3.67]
JN4069	17	1.14	3	17	1.41	3	4.1%	0.00 [-2.05, 2.05]
Total (95% CI)			**71**			**72**	**100.0%**	**-1.11 [-1.69, -0.53]**

Heterogeneity: Tau² = 1.01; Chi² = 60.47, df = 25 (P < 0.0001); I² = 59%
Test for overall effect: Z = 3.74 (P = 0.0002)

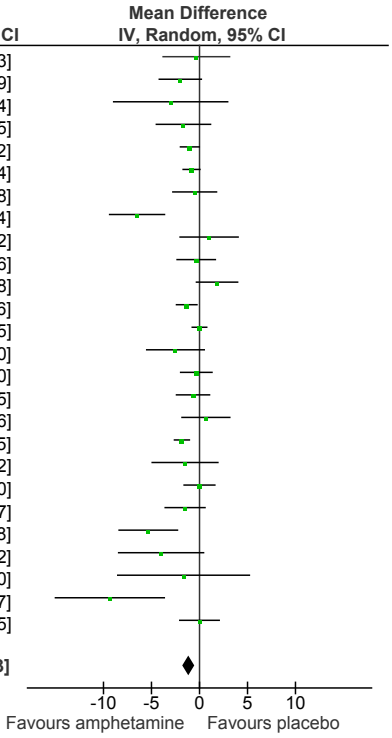


Figure 4.10.2.4 Parent ratings of inattention on the Conners scale

Study or Subgroup	Amphetamine Mean	SD	Total	Placebo Mean	SD	Total	Weight	Mean Difference IV, Random, 95% CI
Duggan1	0.67	0.58	3	13	1	3	4.0%	-12.33 [-13.64, -11.02]
Duggan2	11.67	2.18	3	8.33	5.1	3	3.3%	3.34 [-2.94, 9.62]
JN4000	4	2.16	3	6.33	0.94	3	3.9%	-2.33 [-5.00, 0.34]
JN4005	9.67	0.57	3	10	0.82	3	4.0%	-0.33 [-1.46, 0.80]
JN4006	3	1.63	3	2.67	1.25	3	3.9%	0.33 [-1.99, 2.65]
JN4010	4	4	2	10.5	0.5	2	3.4%	-6.50 [-12.09, -0.91]
JN4011	5.67	4.5	3	7.67	5.44	3	3.0%	-2.00 [-9.99, 5.99]
JN4012	9	2.45	3	3	1.63	3	3.8%	6.00 [2.67, 9.33]
JN4013	4.67	4.11	3	7.33	1.7	3	3.5%	-2.66 [-7.69, 2.37]
JN4014	8.67	4.19	3	7.33	1.7	3	3.5%	1.34 [-3.78, 6.46]
JN4015	2.33	1.89	3	15.67	2.63	3	3.7%	-13.34 [-17.00, -9.68]
JN4020	7.33	1.25	3	7.33	0.94	3	3.9%	0.00 [-1.77, 1.77]
JN4021	0.67	0.58	3	9	2	2	3.8%	-8.33 [-11.18, -5.48]
JN4024	4.33	1.89	3	6.33	2.06	3	3.8%	-2.00 [-5.16, 1.16]
JN4025	4.67	2.36	3	9	2.94	3	3.7%	-4.33 [-8.60, -0.06]
JN4029	10.33	2.49	3	9.33	2.87	3	3.7%	1.00 [-3.30, 5.30]
JN4035	0.33	0.47	3	0.67	0.58	3	4.0%	-0.34 [-1.18, 0.50]
JN4036	0.83	0.58	3	3.67	4.65	3	3.5%	-2.84 [-8.14, 2.46]
JN4037	9	2.78	3	11.5	0.5	3	3.8%	-2.50 [-5.70, 0.70]
JN4039	10.67	4.48	3	8.67	3.25	3	3.3%	2.00 [-4.26, 8.26]
JN4041	5.33	2.25	3	6.17	3.33	3	3.6%	-0.84 [-5.39, 3.71]
JN4042	9.67	2.63	3	7.5	1.73	3	3.8%	2.17 [-1.39, 5.73]
JN4044	5.5	6.06	3	10.17	1.04	3	3.2%	-4.67 [-11.63, 2.29]
JN4048	12.17	1.26	3	12.67	1.26	3	3.9%	-0.50 [-2.52, 1.52]
JN4053	9.33	0.58	3	11	1	3	4.0%	-1.67 [-2.98, -0.36]
JN4063	17.67	0.58	3	17.5	0.71	2	4.0%	0.17 [-1.01, 1.35]
JN4067	0.67	0.58	3	17.67	0.58	3	4.0%	-17.00 [-17.93, -16.07]
Total (95% CI)			**80**			**78**	**100.0%**	**-2.59 [-5.35, 0.18]**

Heterogeneity: Tau² = 49.55; Chi² = 1227.11, df = 26 (P < 0.00001); I² = 98%
Test for overall effect: Z = 1.84 (P = 0.07)

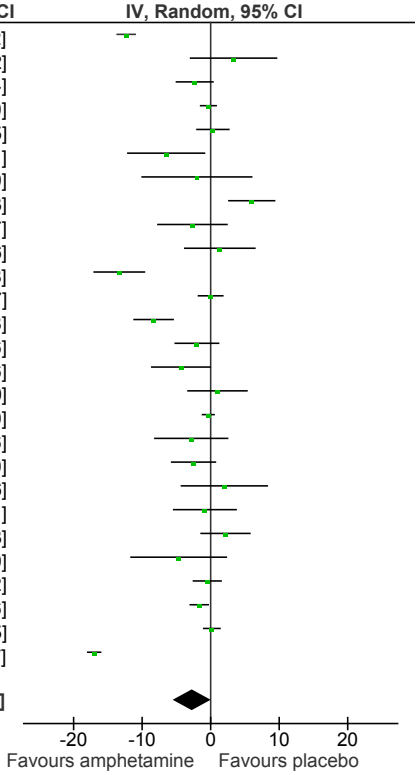


Figure 4.10.2.5 Teacher ratings of hyperactivity/impulsivity on the DuPaul scale

Study or Subgroup	Amphetamine Mean	SD	Total	Placebo Mean	SD	Total	Weight	Mean Difference IV, Random, 95% CI
JN4000	8.5	4.5	2	16.5	5.5	2	2.3%	-8.00 [-17.85, 1.85]
JN4006	4.67	3.86	3	6.67	4.71	3	3.2%	-2.00 [-8.89, 4.89]
JN4011	4.5	0.5	2	6.5	0.5	2	5.3%	-2.00 [-2.98, -1.02]
JN4012	17	1.63	3	15	2	2	4.6%	2.00 [-1.33, 5.33]
JN4014	16.5	0.5	2	8	4	2	3.7%	8.50 [2.91, 14.09]
JN4015	1	1	2	22	2	2	4.7%	-21.00 [-24.10, -17.90]
JN4020	4.33	1.7	3	3.67	2.63	3	4.6%	0.66 [-2.88, 4.20]
JN4024	0.67	0.94	3	7	3.74	3	4.2%	-6.33 [-10.69, -1.97]
JN4025	17	16.67	2	16.5	0.5	2	0.6%	0.50 [-22.61, 23.61]
JN4029	1.33	1.25	3	4.33	1.25	3	5.1%	-3.00 [-5.00, -1.00]
JN4036	1	0.82	3	6.33	3.86	3	4.2%	-5.33 [-9.80, -0.86]
JN4037	1.33	0.58	3	2	3.46	3	4.4%	-0.67 [-4.64, 3.30]
JN4038	1.67	2.08	3	0.67	0.58	3	4.9%	1.00 [-1.44, 3.44]
JN4041	7.33	6.13	3	16.67	0.47	3	3.2%	-9.34 [-16.30, -2.38]
JN4042	6.33	3.3	3	10.67	5.67	3	3.0%	-4.34 [-11.76, 3.08]
JN4045	1.33	0.58	3	1.67	1.55	3	5.1%	-0.34 [-2.21, 1.53]
JN4063	9.33	8.5	3	6	5.2	3	1.9%	3.33 [-7.95, 14.61]
JN4065	1.67	0.47	3	1	1	3	5.2%	0.67 [-0.58, 1.92]
JN4067	7	4.24	3	19	1.41	3	4.0%	-12.00 [-17.06, -6.94]
JN4069	5	6.38	3	10.33	7.32	3	2.0%	-5.33 [-16.32, 5.66]
JN4077	11	5	2	16	2.83	3	3.0%	-5.00 [-12.63, 2.63]
JN4082	6.67	0.94	3	6.67	2.06	3	4.9%	0.00 [-2.56, 2.56]
JN5013	2	1.63	3	10.33	1.7	3	4.9%	-8.33 [-11.00, -5.66]
JN5017	2.33	0.47	3	5	4.08	3	4.1%	-2.67 [-7.32, 1.98]
JN5020	8.33	8.26	3	9.5	7.5	2	1.4%	-1.17 [-15.15, 12.81]
JN5023	1	1	2	5.67	0.94	3	5.1%	-4.67 [-6.42, -2.92]
Total (95% CI)			**71**			**71**	**100.0%**	**-3.38 [-5.35, -1.41]**

Heterogeneity: Tau² = 18.83; Chi² = 254.05, df = 25 (P < 0.00001); I² = 90%
Test for overall effect: Z = 3.37 (P = 0.0008)

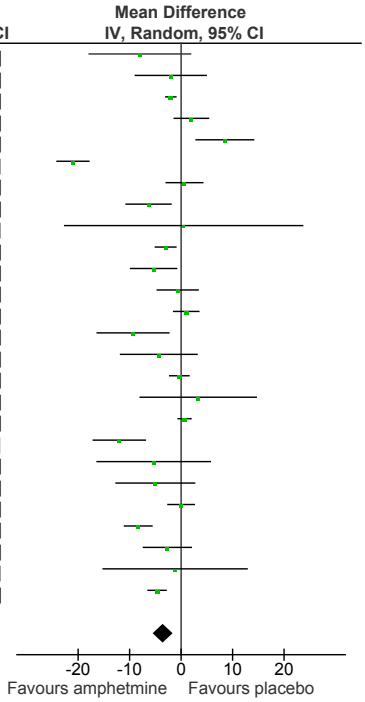


Figure 4.10.2.6 Parent ratings of hyperactivity/impulsivity on the DuPaul scale

Study or Subgroup	Amphetamine Mean	SD	Total	Placebo Mean	SD	Total	Weight	Mean Difference IV, Random, 95% CI
JN4000	18.33	4.11	3	18.67	0.47	3	3.3%	-0.34 [-5.02, 4.34]
JN4005	8	0.82	3	9.33	0.47	3	4.2%	-1.33 [-2.40, -0.26]
JN4006	8	2.83	3	10.67	3.09	3	3.2%	-2.67 [-7.41, 2.07]
JN4010	6	4	2	9	2	2	2.8%	-3.00 [-9.20, 3.20]
JN4011	2.33	1.25	3	2.67	1.25	3	4.0%	-0.34 [-2.34, 1.66]
JN4012	4.67	2.49	3	2.67	0.94	3	3.8%	2.00 [-1.01, 5.01]
JN4013	5	2.94	3	12.33	4.03	3	3.0%	-7.33 [-12.97, -1.69]
JN4014	6	2.45	3	7	3.27	3	3.3%	-1.00 [-5.62, 3.62]
JN4015	5.67	2.63	3	23.67	0.47	3	3.8%	-18.00 [-21.02, -14.98]
JN4020	1	0.82	3	0.33	0.47	3	4.2%	0.67 [-0.40, 1.74]
JN4021	2.33	1.89	3	17	1.63	3	3.8%	-14.67 [-17.49, -11.85]
JN4024	7.67	1.7	3	8.33	2.06	3	3.8%	-0.66 [-3.68, 2.36]
JN4025	16.67	1.25	3	21	1.63	3	3.9%	-4.33 [-6.65, -2.01]
JN4029	8	0.82	3	9.67	0.94	3	4.1%	-1.67 [-3.08, -0.26]
JN4035	3.33	2.63	3	0.33	0.47	3	3.8%	3.00 [-0.02, 6.02]
JN4036	5	4.55	3	8	4.24	3	2.5%	-3.00 [-10.04, 4.04]
JN4037	13.33	1.15	3	20.67	2.08	3	3.8%	-7.34 [-10.03, -4.65]
JN4039	14.67	5.31	3	13.33	4.92	3	2.2%	1.34 [-6.85, 9.53]
JN4041	1.33	0.47	3	9.33	6.94	3	2.3%	-8.00 [-15.87, -0.13]
JN4042	8	5.35	3	7.33	2.06	3	2.7%	0.67 [-5.82, 7.16]
JN4045	7	1.73	3	12.67	6.03	3	2.5%	-5.67 [-12.77, 1.43]
JN4063	8.67	5.51	3	13	7.07	2	1.5%	-4.33 [-15.94, 7.28]
JN4065	0.33	0.47	3	1	1.73	3	4.0%	-0.67 [-2.70, 1.36]
JN4067	2.67	2.06	3	18.67	1.25	3	3.8%	-16.00 [-18.73, -13.27]
JN4069	11	1.41	3	12.67	3.09	3	3.5%	-1.67 [-5.51, 2.17]
JN4077	16	1	2	16	5.35	3	2.8%	0.00 [-6.21, 6.21]
JN4082	2	0.82	3	2.33	1.25	3	4.1%	-0.33 [-2.02, 1.36]
JN5013	4	0.82	3	5.67	0.94	3	4.1%	-1.67 [-3.08, -0.26]
JN5020	18.67	1.7	3	15	5.89	3	2.6%	3.67 [-3.27, 10.61]
JN5023	14	1	3	16	5.1	3	2.9%	-2.00 [-7.88, 3.88]
Total (95% CI)			**88**			**88**	**100.0%**	**-3.23 [-5.00, -1.46]**

Heterogeneity: Tau² = 19.33; Chi² = 368.81, df = 29 (P < 0.00001); I² = 92%
Test for overall effect: Z = 3.57 (P = 0.0004)

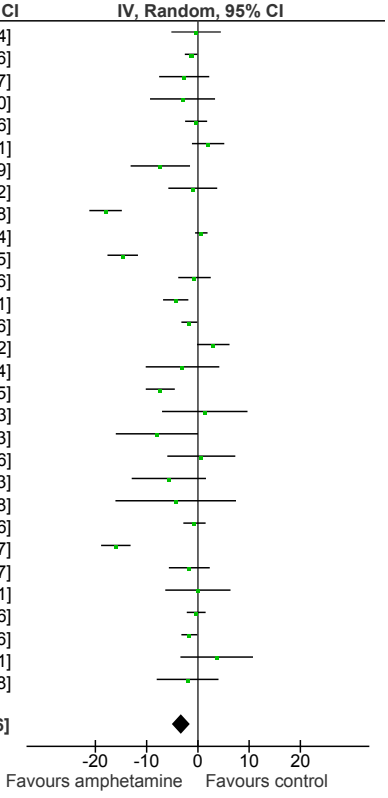


Figure 4.10.2.7 Teacher ratings of hyperactivity on the Conners scale

Study or Subgroup	Amphetamine Mean	SD	Total	Placebo Mean	SD	Total	Weight	Mean Difference IV, Random, 95% CI
Duggan1	1.33	2.31	3	0.67	1.15	3	5.1%	0.66 [-2.26, 3.58]
Duggan2	1.67	0.58	3	2.67	1.15	3	6.4%	-1.00 [-2.46, 0.46]
JN4000	9	7	2	17.5	1.5	2	1.3%	-8.50 [-18.42, 1.42]
JN4005	14.5	3.5	2	18.33	3.77	3	2.5%	-3.83 [-10.29, 2.63]
JN4006	3.33	2.06	3	4	2.94	3	4.0%	-0.67 [-4.73, 3.39]
JN4011	3.5	0.5	2	4.5	0.5	2	6.7%	-1.00 [-1.98, -0.02]
JN4012	9.33	1.25	3	10.5	0.5	2	6.3%	-1.17 [-2.75, 0.41]
JN4014	2.5	0.5	2	3.67	1.7	3	5.9%	-1.17 [-3.21, 0.87]
JN4015	0.5	0.5	2	20.5	3.5	2	3.4%	-20.00 [-24.90, -15.10]
JN4020	1.67	1.25	3	3	2.45	3	4.9%	-1.33 [-4.44, 1.78]
JN4025	11	1	2	10.33	2.49	3	4.9%	0.67 [-2.47, 3.81]
JN4029	0.33	0.47	3	3.33	2.06	3	5.6%	-3.00 [-5.39, -0.61]
JN4037	1.33	0.58	3	2	2.18	3	5.4%	-0.67 [-3.22, 1.88]
JN4038	1	1	3	1	1	3	6.3%	0.00 [-1.60, 1.60]
JN4039	8.5	4.5	3	3.75	2.47	2	2.6%	4.75 [-1.39, 10.89]
JN4042	4.67	3.75	3	8.17	2.75	3	3.1%	-3.50 [-8.76, 1.76]
JN4044	3.33	1.25	3	6.83	2.75	3	4.6%	-3.50 [-6.92, -0.08]
JN4048	8.25	0.35	2	9.5	5.66	2	1.9%	-1.25 [-9.11, 6.61]
JN4053	3.33	0.58	3	8.67	2.89	3	4.7%	-5.34 [-8.68, -2.00]
JN4063	14	5	3	15.33	8.14	3	1.1%	-1.33 [-12.14, 9.48]
JN4064	0.67	0.58	3	2	2	3	5.6%	-1.33 [-3.69, 1.03]
JN4067	6.67	4.71	3	20.67	0.47	3	3.1%	-14.00 [-19.36, -8.64]
JN4069	15.33	2.49	3	17.33	0.94	3	5.0%	-2.00 [-5.01, 1.01]
Total (95% CI)			**62**			**63**	**100.0%**	**-2.42 [-3.68, -1.16]**

Heterogeneity: Tau² = 5.90; Chi² = 100.26, df = 22 (P < 0.00001); I² = 78%
Test for overall effect: Z = 3.78 (P = 0.0002)

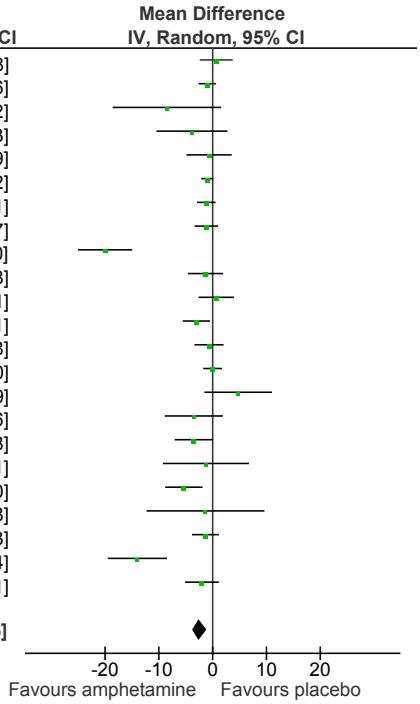


Figure 4.10.2.8 Parent ratings of hyperactivity on the Conners scale

Study or Subgroup	Amphetamine Mean	SD	Total	Placebo Mean	SD	Total	Weight	Mean Difference IV, Random, 95% CI
Duggan1	3.67	0.58	3	0.67	0.58	3	4.3%	3.00 [2.07, 3.93]
Duggan2	2.33	2.52	3	1.33	2.31	3	3.8%	1.00 [-2.87, 4.87]
JN4000	12	3.74	3	15	1.63	3	3.6%	-3.00 [-7.62, 1.62]
JN4005	7.67	0.47	3	9.33	1.89	3	4.1%	-1.66 [-3.86, 0.54]
JN4006	6	2.94	3	5.33	1.25	3	3.8%	0.67 [-2.95, 4.29]
JN4010	3.5	3.5	2	5.5	0.5	2	3.5%	-2.00 [-6.90, 2.90]
JN4011	1.33	1.25	3	1	0.82	3	4.2%	0.33 [-1.36, 2.02]
JN4012	3.33	2.63	3	1.33	0.47	3	4.0%	2.00 [-1.02, 5.02]
JN4013	4.33	4.19	3	8.33	3.3	3	3.2%	-4.00 [-10.04, 2.04]
JN4014	5.33	0.47	3	4.67	0.47	3	4.3%	0.66 [-0.09, 1.41]
JN4015	3.67	1.7	3	17.67	0.58	3	4.2%	-14.00 [-16.03, -11.97]
JN4021	2.33	0.47	3	11.67	0.58	2	4.3%	-9.34 [-10.30, -8.38]
JN4024	3.67	1.7	3	5.67	0.94	3	4.1%	-2.00 [-4.20, 0.20]
JN4025	8.67	3.09	3	12.33	1.7	3	3.7%	-3.66 [-7.65, 0.33]
JN4029	5	1.63	3	5.67	1.89	3	4.0%	-0.67 [-3.49, 2.15]
JN4035	0.33	0.47	3	0.67	0.58	3	4.3%	-0.34 [-1.18, 0.50]
JN4036	0.83	1.04	3	2.33	2.84	3	3.9%	-1.50 [-4.92, 1.92]
JN4037	11.83	0.29	3	17	0.87	3	4.3%	-5.17 [-6.21, -4.13]
JN4039	10.67	4.8	3	8.67	2.75	3	3.1%	2.00 [-4.26, 8.26]
JN4041	1.83	0.76	3	3.5	3	3	3.9%	-1.67 [-5.17, 1.83]
JN4042	8	2.83	3	4.83	3.62	3	3.4%	3.17 [-2.03, 8.37]
JN4044	3.17	3.88	3	5.83	0.29	3	3.6%	-2.66 [-7.06, 1.74]
JN4048	13.5	3	3	14.67	1.53	3	3.8%	-1.17 [-4.98, 2.64]
JN4053	9.33	2.08	3	10	2	3	3.9%	-0.67 [-3.94, 2.60]
JN4063	11.33	6.43	3	16	2.83	2	2.5%	-4.67 [-12.94, 3.60]
JN4067	1	1.41	3	14.67	0.58	3	4.2%	-13.67 [-15.40, -11.94]
Total (95% CI)			**77**			**75**	**100.0%**	**-2.34 [-4.39, -0.29]**

Heterogeneity: Tau² = 25.20; Chi² = 766.89, df = 25 (P < 0.00001); I² = 97%
Test for overall effect: Z = 2.24 (P = 0.03)

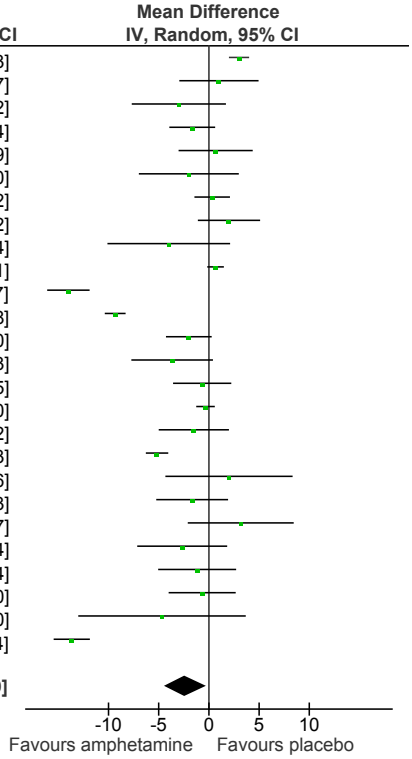


Figure 4.10.2.9 Teacher ratings of ADHD index on the Conners scale

Study or Subgroup	Amphetamine Mean	SD	Total	Placebo Mean	SD	Total	Weight	Mean Difference IV, Random, 95% CI
Duggan1	3.33	5.77	3	5.33	3.21	3	4.0%	-2.00 [-9.47, 5.47]
Duggan2	3.67	5.51	3	8.33	1.53	3	4.2%	-4.66 [-11.13, 1.81]
JN4005	26	9	2	27.33	6.6	3	2.5%	-1.33 [-15.87, 13.21]
JN4006	11.67	4.5	3	16.67	4.78	3	4.0%	-5.00 [-12.43, 2.43]
JN4011	9	2	2	13.5	0.5	2	4.9%	-4.50 [-7.36, -1.64]
JN4012	32.33	0.47	3	31.5	1.5	2	5.0%	0.83 [-1.32, 2.98]
JN4014	16.5	0.5	2	16.33	2.87	3	4.8%	0.17 [-3.15, 3.49]
JN4015	2	2	2	34	2	2	4.7%	-32.00 [-35.92, -28.08]
JN4020	14.33	5.74	3	14.33	5.44	3	3.6%	0.00 [-8.95, 8.95]
JN4021	0.33	0.47	3	3.33	2.36	3	4.9%	-3.00 [-5.72, -0.28]
JN4024	0.33	0.47	3	12.67	5.76	3	4.2%	-12.34 [-18.88, -5.80]
JN4025	25.5	0.5	2	25.33	2.06	3	4.9%	0.17 [-2.26, 2.60]
JN4029	6	3.56	3	12.67	4.92	3	4.1%	-6.67 [-13.54, 0.20]
JN4036	1.67	0.47	3	8	5.77	3	4.2%	-6.33 [-12.88, 0.22]
JN4037	6.67	0.47	3	5	5.29	3	4.3%	1.67 [-4.34, 7.68]
JN4038	3.33	2.31	3	3	1	3	4.9%	0.33 [-2.52, 3.18]
JN4039	16.5	8.05	3	10	4.95	2	3.1%	6.50 [-4.90, 17.90]
JN4042	12.5	2.29	3	17.83	3.18	3	4.6%	-5.33 [-9.76, -0.90]
JN4044	9.17	3.21	3	15.17	4.19	3	4.3%	-6.00 [-11.97, -0.03]
JN4048	16	3.54	2	17.75	10.25	2	2.4%	-1.75 [-16.78, 13.28]
JN4053	3.67	0.47	3	11.33	5.86	3	4.1%	-7.66 [-14.31, -1.01]
JN4063	9.67	11.59	3	20.67	12.86	3	1.7%	-11.00 [-30.59, 8.59]
JN4064	7.67	8.62	3	11.67	4.62	3	3.1%	-4.00 [-15.07, 7.07]
JN4067	15.33	11.32	3	35.67	0.47	3	2.8%	-20.34 [-33.16, -7.52]
JN4069	34.67	1.89	3	35	1.41	3	4.9%	-0.33 [-3.00, 2.34]
Total (95% CI)			**69**			**70**	**100.0%**	**-4.85 [-8.01, -1.68]**

Heterogeneity: Tau² = 51.42; Chi² = 270.13, df = 24 (P < 0.00001); I² = 91%
Test for overall effect: Z = 3.00 (P = 0.003)

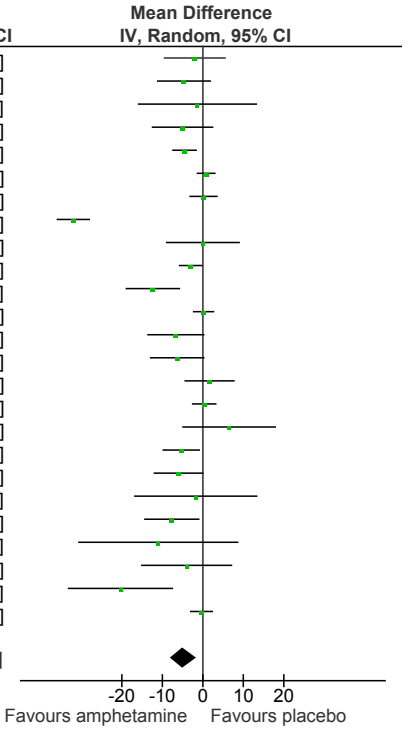


Figure 4.10.2.10 Parent ratings of ADHD index on the Conners scale

Study or Subgroup	Amphetamine Mean	SD	Total	Placebo Mean	SD	Total	Weight	Mean Difference IV, Random, 95% CI
Duggan1	3.33	0.58	3	27	1	3	4.2%	-23.67 [-24.98, -22.36]
Duggan2	19.33	4.8	3	13.67	4.27	3	3.8%	5.66 [-1.61, 12.93]
JN4000	20.67	6.65	3	28	1.63	3	3.8%	-7.33 [-15.08, 0.42]
JN4005	18.33	0.47	3	19.67	0.94	3	4.2%	-1.34 [-2.53, -0.15]
JN4006	13.33	8.5	3	12.67	3.68	3	3.5%	0.66 [-9.82, 11.14]
JN4010	10	10	2	20.5	0.5	2	3.1%	-10.50 [-24.38, 3.38]
JN4011	10.33	6.55	3	9	5.72	3	3.6%	1.33 [-8.51, 11.17]
JN4012	16.33	3.68	3	11	1.63	3	4.1%	5.33 [0.78, 9.88]
JN4013	13.33	6.6	3	20.33	4.03	3	3.7%	-7.00 [-15.75, 1.75]
JN4014	15.67	4.03	3	16.67	0.47	3	4.1%	-1.00 [-5.59, 3.59]
JN4015	7.67	3.77	3	34.67	1.25	3	4.1%	-27.00 [-31.49, -22.51]
JN4020	10	0.82	3	11.33	3.3	3	4.1%	-1.33 [-5.18, 2.52]
JN4021	3.67	1.25	3	20	2	2	4.2%	-16.33 [-19.44, -13.22]
JN4024	17.33	7.59	3	22.33	3.3	3	3.6%	-5.00 [-14.37, 4.37]
JN4025	19.67	5.25	3	26.33	2.36	3	3.9%	-6.66 [-13.17, -0.15]
JN4029	20.67	5.56	3	22.33	3.4	3	3.8%	-1.66 [-9.03, 5.71]
JN4036	7	4.33	3	10.83	10.56	3	3.2%	-3.83 [-16.75, 9.09]
JN4037	21.33	1.89	3	29.67	0.58	3	4.2%	-8.34 [-10.58, -6.10]
JN4039	24	8.72	3	19.17	5.48	3	3.4%	4.83 [-6.82, 16.48]
JN4041	9	2.65	3	10.67	5.01	3	3.9%	-1.67 [-8.08, 4.74]
JN4042	19.83	3.4	3	14.83	3.69	3	4.0%	5.00 [-0.68, 10.68]
JN4044	9.5	10.97	3	17.5	1.32	3	3.3%	-8.00 [-20.50, 4.50]
JN4048	24.5	2.5	3	27.67	6.45	3	3.8%	-3.17 [-11.00, 4.66]
JN4053	19	1	3	18.67	1.15	3	4.2%	0.33 [-1.39, 2.05]
JN4063	34.33	2.08	3	34.5	2.12	2	4.1%	-0.17 [-3.93, 3.59]
JN4067	3	4.24	3	35.67	0.58	3	4.1%	-32.67 [-37.51, -27.83]
Total (95% CI)			**77**			**75**	**100.0%**	**-5.67 [-10.33, -1.01]**

Heterogeneity: Tau² = 133.18; Chi² = 1078.66, df = 25 (P < 0.00001); I² = 98%
Test for overall effect: Z = 2.39 (P = 0.02)

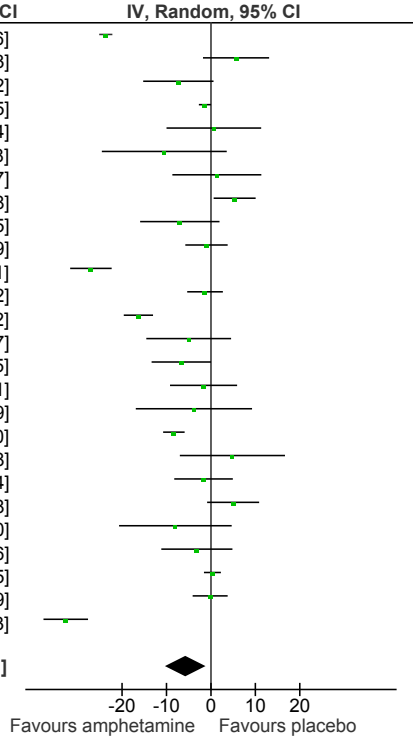


Figure 4.10.3 Flow of studies (methylphenidate)

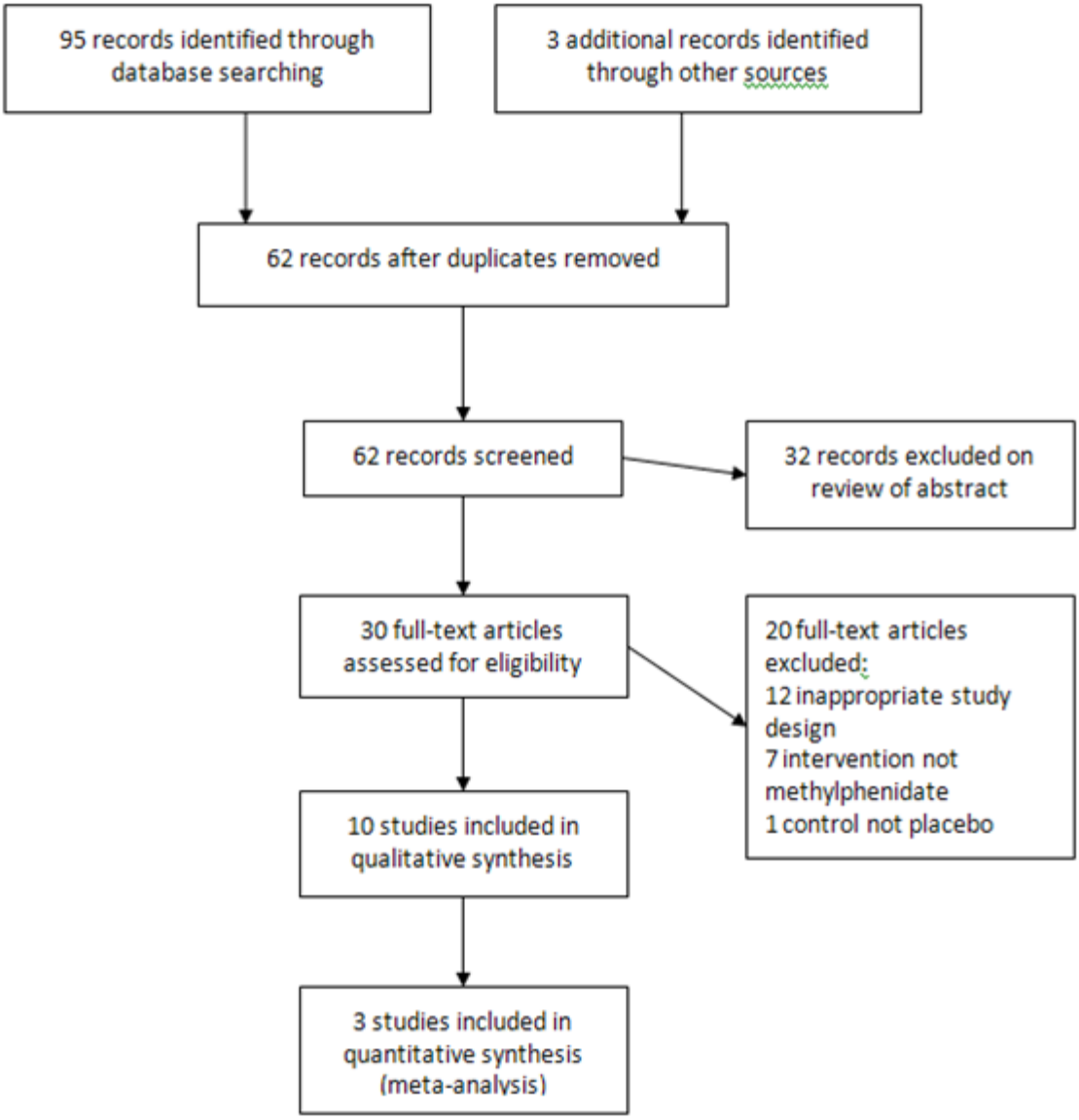


Figure 4.10.4: Meta-analysis (methylphenidate)

Figure 4.10.4.1 Teacher ratings of inattention on the DuPaul scale

Study or Subgroup	Methylphenidate Mean	SD	Total	Placebo Mean	SD	Total	Weight	Mean Difference IV, Random, 95% CI
JN4001	8.5	1.5	2	9.33	2.49	3	3.4%	-0.83 [-4.33, 2.67]
JN4004	7	3	2	21.5	1.5	2	3.2%	-14.50 [-19.15, -9.85]
JN4008	1.33	0.94	3	9.67	0.43	3	3.7%	-8.34 [-9.51, -7.17]
JN4018	4.33	3.3	3	10	4.97	3	2.7%	-5.67 [-12.42, 1.08]
JN4019	9	4.24	3	11	1.41	3	3.1%	-2.00 [-7.06, 3.06]
JN4023	10.67	2.87	3	7.33	0.47	3	3.4%	3.34 [0.05, 6.63]
JN4026	8	1.28	3	5.33	1.25	3	3.6%	2.67 [0.65, 4.69]
JN4027	8.33	2.87	3	20.33	5.25	3	2.7%	-12.00 [-18.77, -5.23]
JN4032	7	2.94	3	17.33	2.05	3	3.3%	-10.33 [-14.39, -6.27]
JN4033	11	3.74	3	19.33	2.05	3	3.1%	-8.33 [-13.16, -3.50]
JN4034	29	3	2	26	2	2	3.1%	3.00 [-2.00, 8.00]
JN4046	0.33	0.47	3	2.5	0.5	2	3.7%	-2.17 [-3.04, -1.30]
JN4061a	3.5	0.5	2	4	1.41	3	3.7%	-0.50 [-2.24, 1.24]
JN4066	14	2.16	3	7	2.45	3	3.4%	7.00 [3.30, 10.70]
JN4072	3	0.82	3	8	2	2	3.5%	-5.00 [-7.92, -2.08]
JN4073	1	1	2	24	3	2	3.2%	-23.00 [-27.38, -18.62]
JN4078	2.33	1.89	3	23.67	1.25	3	3.6%	-21.34 [-23.90, -18.78]
JN4080	15.67	6.6	3	17.67	0.47	3	2.6%	-2.00 [-9.49, 5.49]
JN5001	9	2.94	3	7.67	0.47	3	3.4%	1.33 [-2.04, 4.70]
JN5002	4	1.41	3	3.67	0.94	3	3.6%	0.33 [-1.59, 2.25]
JN5003	9.67	2.62	3	11.33	0.94	3	3.5%	-1.66 [-4.81, 1.49]
JN5005	0.67	0.94	3	1.33	1.25	3	3.7%	-0.66 [-2.43, 1.11]
JN5006	4	3.74	3	17.5	1.5	2	3.2%	-13.50 [-18.22, -8.78]
JN5010	17.67	3.3	3	20.67	3.09	3	3.1%	-3.00 [-8.12, 2.12]
JN5011	14	1.4	3	23.67	1.25	3	3.6%	-9.67 [-11.79, -7.55]
JN5012	15	2	2	13.67	1.25	3	3.5%	1.33 [-1.78, 4.44]
JN5014	9.5	0.5	2	9.33	0.47	3	3.7%	0.17 [-0.70, 1.04]
JN5018	3.33	1.89	3	3.33	3.4	3	3.2%	0.00 [-4.40, 4.40]
JN5026	3.67	3.3	3	18.33	0.94	3	3.3%	-14.66 [-18.54, -10.78]
JN5027	10.33	2.87	3	14	2	2	3.3%	-3.67 [-7.94, 0.60]
Total (95% CI)			**83**			**83**	**100.0%**	**-4.67 [-6.79, -2.56]**

Heterogeneity: Tau² = 31.01; Chi² = 674.47, df = 29 (P < 0.00001); I² = 96%
Test for overall effect: Z = 4.33 (P < 0.0001)

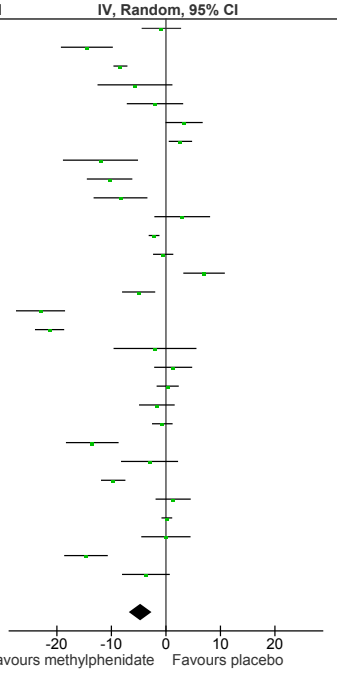


Figure 4.10.4.2 Parent ratings of inattention on the DuPaul scale

Study or Subgroup	Methylphenidate Mean	SD	Total	Placebo Mean	SD	Total	Weight	Mean Difference IV, Random, 95% CI
JN4001	7	1.41	3	14.67	0.47	3	3.8%	-7.67 [-9.35, -5.99]
JN4004	17.33	0.47	3	21.67	0.47	3	3.9%	-4.34 [-5.09, -3.59]
JN4008	6.67	2.36	3	14.67	0.47	3	3.6%	-8.00 [-10.72, -5.28]
JN4018	17	2	2	19.33	3.4	3	3.0%	-2.33 [-7.07, 2.41]
JN4019	12	0.5	2	15	0.82	3	3.9%	-3.00 [-4.16, -1.84]
JN4023	4	1.63	3	3.67	2.48	3	3.4%	0.33 [-3.03, 3.69]
JN4027	3.5	1.5	2	10	4.24	3	2.9%	-6.50 [-11.73, -1.27]
JN4033	16.5	0.5	2	17.67	3.3	3	3.3%	-1.17 [-4.97, 2.63]
JN4034	20.33	4.78	3	28	3.74	3	2.4%	-7.67 [-14.54, -0.80]
JN4043	11.67	2.87	3	18.33	2.36	3	3.2%	-6.66 [-10.86, -2.46]
JN4060	10	2.94	3	13.33	1.7	3	3.3%	-3.33 [-7.17, 0.51]
JN4066	10.33	2.62	3	7	2.16	3	3.3%	3.33 [-0.51, 7.17]
JN4072	6.67	2.87	3	22	0.82	3	3.4%	-15.33 [-18.71, -11.95]
JN4073	4.5	0.5	2	22.5	4.5	2	2.6%	-18.00 [-24.27, -11.73]
JN4078	16.67	2.62	3	14.67	2.49	3	3.2%	2.00 [-2.09, 6.09]
JN4080	15	3.74	3	15.33	1.7	3	3.0%	-0.33 [-4.98, 4.32]
JN5001	5.67	0.47	3	3	1.41	3	3.8%	2.67 [0.99, 4.35]
JN5002	15.5	2.5	2	16.5	3.5	2	2.7%	-1.00 [-6.96, 4.96]
JN5003	19.67	1.25	3	19.33	1.7	3	3.7%	0.34 [-2.05, 2.73]
JN5004	21	0.82	3	24.33	3.3	3	3.3%	-3.33 [-7.18, 0.52]
JN5005	1	0.82	3	11.33	4.92	3	2.7%	-10.33 [-15.97, -4.69]
JN5006	11.67	4.5	3	8.33	2.49	3	2.7%	3.34 [-2.48, 9.16]
JN5010	20.67	0.47	3	26.67	0.47	3	3.9%	-6.00 [-6.75, -5.25]
JN5011	5.67	0.47	3	7.33	1.25	3	3.8%	-1.66 [-3.17, -0.15]
JN5012	23.67	0.47	3	19.67	0.47	3	3.9%	4.00 [3.25, 4.75]
JN5014	13	4.55	3	11.67	1.89	3	2.8%	1.33 [-4.25, 6.91]
JN5015	10.67	1.25	3	10	0.82	3	3.8%	0.67 [-1.02, 2.36]
JN5018	11	2.83	3	12.33	1.7	3	3.3%	-1.33 [-5.07, 2.41]
JN5026	9.33	0.47	3	10	3.1	3	3.4%	-0.67 [-4.22, 2.88]
JN5027	8.33	0.47	3	6.67	0.47	3	3.9%	1.66 [0.91, 2.41]
Total (95% CI)			**84**			**88**	**100.0%**	**-2.78 [-4.47, -1.08]**

Heterogeneity: Tau² = 18.87; Chi² = 706.20, df = 29 (P < 0.00001); I² = 96%
Test for overall effect: Z = 3.21 (P = 0.001)

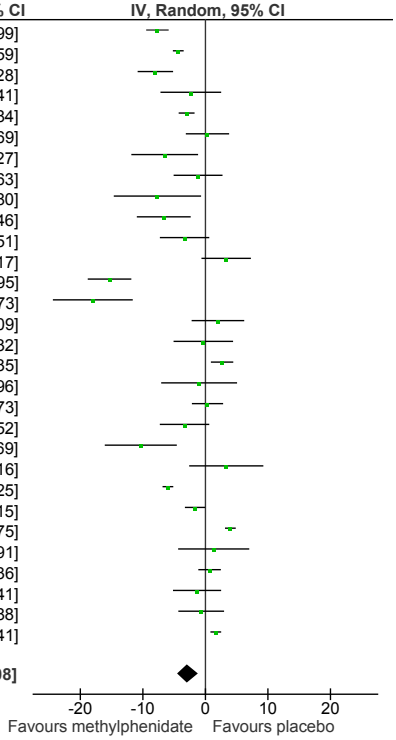


Figure 4.10.4.3 Teacher ratings of inattention on the Conners scale

Study or Subgroup	Methylphenidate Mean	SD	Total	Placebo Mean	SD	Total	Weight	Mean Difference IV, Random, 95% CI
JN4001	4.33	0.94	3	12.67	1.89	3	8.2%	-8.34 [-10.73, -5.95]
JN4004	1.33	1.89	3	2.33	1.25	3	7.9%	-1.00 [-3.56, 1.56]
JN4008	0.33	0.47	3	0.67	0.94	3	10.0%	-0.34 [-1.53, 0.85]
JN4018	5.33	1.89	3	6.33	2.49	3	6.4%	-1.00 [-4.54, 2.54]
JN4019	7.67	0.47	3	8.67	1.89	3	8.5%	-1.00 [-3.20, 1.20]
JN4023	3	1.41	3	0.33	0.47	3	9.4%	2.67 [0.99, 4.35]
JN4026	0.67	0.47	3	0.67	0.94	3	10.0%	0.00 [-1.19, 1.19]
JN4027	9.67	1.7	3	12.67	2.62	3	6.4%	-3.00 [-6.53, 0.53]
JN4032	5.678	1.7	3	9	2.83	3	6.1%	-3.32 [-7.06, 0.41]
JN4033	5	2.94	3	8.33	1.7	3	6.0%	-3.33 [-7.17, 0.51]
JN4034	9	1.41	3	9.67	2.87	3	6.3%	-0.67 [-4.29, 2.95]
JN4046	1.33	0.47	3	1	0.82	3	10.2%	0.33 [-0.74, 1.40]
JN4049	3	2.16	3	7.33	3.86	3	4.5%	-4.33 [-9.34, 0.68]
Total (95% CI)			**39**			**39**	**100.0%**	**-1.50 [-2.90, -0.10]**

Heterogeneity: Tau² = 4.72; Chi² = 66.97, df = 12 (P < 0.00001); I² = 82%
Test for overall effect: Z = 2.10 (P = 0.04)

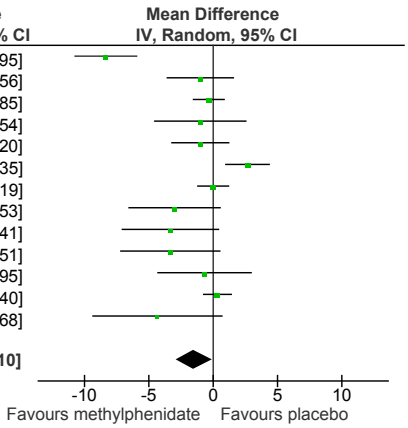


Figure 4.10.4.4 Parent ratings of inattention on the Conners scale

Study or Subgroup	Methylphenidate Mean	SD	Total	Placebo Mean	SD	Total	Weight	Mean Difference IV, Random, 95% CI
JN4001	5	2	2	11.5	0.5	2	6.6%	-6.50 [-9.36, -3.64]
JN4019	10	0.82	3	13	2.16	3	7.1%	-3.00 [-5.61, -0.39]
JN4023	1.33	1.25	3	1	0.82	3	9.4%	0.33 [-1.36, 2.02]
JN4043	9	1.41	3	10.33	1.7	3	7.4%	-1.33 [-3.83, 1.17]
JN4046	8.5	0.5	2	6.5	0.5	2	11.0%	2.00 [1.02, 2.98]
JN4054	0.5	0.5	2	0.67	0.47	3	11.2%	-0.17 [-1.04, 0.70]
JN5004	4	0.82	3	5	1.41	3	9.0%	-1.00 [-2.85, 0.85]
JN5005	2	1.63	3	4.33	0.47	3	8.8%	-2.33 [-4.25, -0.41]
JN5006	4.33	1.7	3	7	2.94	3	4.8%	-2.67 [-6.51, 1.17]
JN5015	16	0.82	3	14.67	1.25	3	9.4%	1.33 [-0.36, 3.02]
JN5018	8.67	3.77	3	9.67	1.25	3	4.0%	-1.00 [-5.49, 3.49]
JN5027	4.33	0.47	3	3.33	0.47	3	11.4%	1.00 [0.25, 1.75]
Total (95% CI)			**33**			**34**	**100.0%**	**-0.73 [-1.82, 0.36]**

Heterogeneity: Tau² = 2.56; Chi² = 58.65, df = 11 (P < 0.00001); I² = 81%
Test for overall effect: Z = 1.32 (P = 0.19)

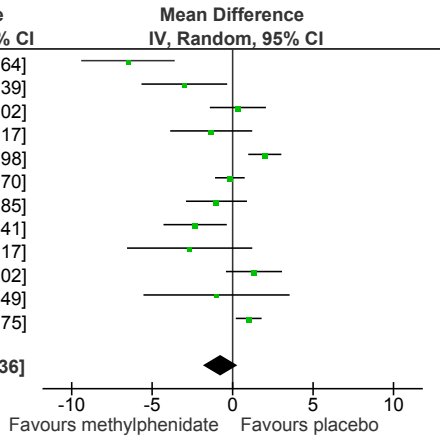


Figure 4.10.4.5 Teacher ratings of hyperactivity/impulsivity on the DuPaul scale

Study or Subgroup	Methylphenidate Mean	SD	Total	Placebo Mean	SD	Total	Weight	Mean Difference IV, Random, 95% CI
JN4001	6	1	2	6	3.27	3	3.3%	0.00 [-3.95, 3.95]
JN4004	7	1	2	12.5	0.5	2	3.8%	-5.50 [-7.05, -3.95]
JN4008	1.33	1.25	3	8.33	2.05	3	3.6%	-7.00 [-9.72, -4.28]
JN4018	0.33	0.47	3	1	1.41	3	3.8%	-0.67 [-2.35, 1.01]
JN4019	0.33	0.47	3	0.67	0.94	3	3.9%	-0.34 [-1.53, 0.85]
JN4023	10	3.27	3	5.33	4.11	3	2.7%	4.67 [-1.27, 10.61]
JN4026	5.33	1.89	3	3.33	2.62	3	3.4%	2.00 [-1.66, 5.66]
JN4027	3	2.16	3	17	5.72	3	2.5%	-14.00 [-20.92, -7.08]
JN4032	4	1.41	3	13.33	6.02	3	2.4%	-9.33 [-16.33, -2.33]
JN4033	0.33	0.47	3	8	4.55	3	2.9%	-7.67 [-12.85, -2.49]
JN4034	15.5	2.5	2	10.5	0.5	2	3.4%	5.00 [1.47, 8.53]
JN4046	1.33	1.89	3	1	1	2	3.6%	0.33 [-2.22, 2.88]
JN4061a	6.5	0.5	2	6	2.16	3	3.6%	0.50 [-2.04, 3.04]
JN4066	12.67	3.86	3	14.67	0.94	3	3.1%	-2.00 [-6.50, 2.50]
JN4072	4.66	0.94	3	10	4	2	2.8%	-5.34 [-10.98, 0.30]
JN4073	3.5	1.5	2	25	2	2	3.4%	-21.50 [-24.96, -18.04]
JN4078	2.67	3.09	3	25.67	0.47	3	3.4%	-23.00 [-26.54, -19.46]
JN4080	1.33	1.89	3	4	3.27	3	3.2%	-2.67 [-6.94, 1.60]
JN5001	5.33	3.09	3	5	1.63	3	3.3%	0.33 [-3.62, 4.28]
JN5002	2.33	0.47	3	3	1.41	3	3.8%	-0.67 [-2.35, 1.01]
JN5003	9.67	3.3	3	11.67	0.47	3	3.3%	-2.00 [-5.77, 1.77]
JN5005	0.67	0.94	3	2.33	1.7	3	3.7%	-1.66 [-3.86, 0.54]
JN5006	6.67	3.68	3	9.5	0.5	2	3.2%	-2.83 [-7.05, 1.39]
JN5010	10	2.32	3	10.67	2.19	3	3.4%	-0.67 [-4.28, 2.94]
JN5011	8.33	0.94	3	6.67	2.05	3	3.6%	1.66 [-0.89, 4.21]
JN5012	18	2	1	15	3.07	3	2.9%	3.00 [-2.24, 8.24]
JN5014	8	1	2	6.33	0.47	3	3.8%	1.67 [0.19, 3.15]
JN5018	0.33	0.47	3	0.33	0.47	3	3.9%	0.00 [-0.75, 0.75]
JN5026	2	2.16	3	14.33	3.86	3	3.0%	-12.33 [-17.34, -7.32]
JN5027	5	2.45	3	13	2	2	3.3%	-8.00 [-11.92, -4.08]
Total (95% CI)			**82**			**83**	**100.0%**	**-3.39 [-5.16, -1.62]**

Heterogeneity: Tau² = 20.78; Chi² = 438.82, df = 29 (P < 0.00001); I² = 93%
Test for overall effect: Z = 3.75 (P = 0.0002)

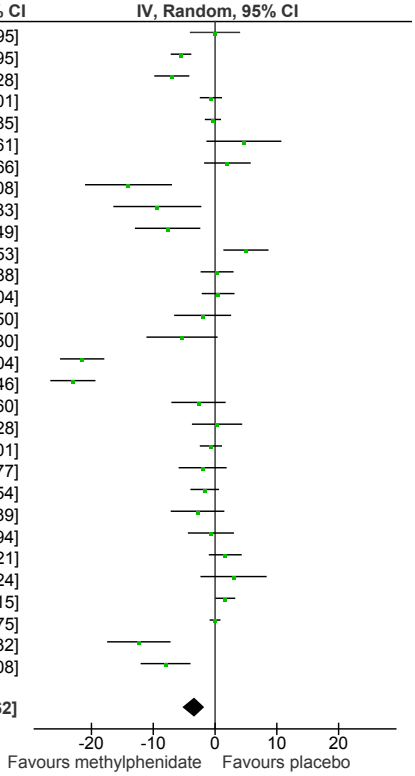


Figure 4.10.4.6 Parent ratings of hyperactivity/impulsivity on the DuPaul scale

Study or Subgroup	Methylphenidate Mean	SD	Total	Placebo Mean	SD	Total	Weight	Mean Difference IV, Random, 95% CI
JN4001	1.67	0.47	3	6.67	1.25	3	4.0%	-5.00 [-6.51, -3.49]
JN4004	14	0.48	3	21	1.41	3	3.9%	-7.00 [-8.69, -5.31]
JN4008	7.67	0.47	3	11.33	2.05	3	3.7%	-3.66 [-6.04, -1.28]
JN4018	7.5	3.5	2	12.67	1.89	3	2.5%	-5.17 [-10.47, 0.13]
JN4019	18	2	3	17.33	6.65	3	1.7%	0.67 [-7.19, 8.53]
JN4023	2.67	0.94	3	5.33	0.94	3	4.0%	-2.66 [-4.16, -1.16]
JN4027	5	1	2	12	4.32	3	2.6%	-7.00 [-12.08, -1.92]
JN4033	10	3	2	12	1.63	3	2.8%	-2.00 [-6.55, 2.55]
JN4034	11.33	2.05	3	18.67	4.19	3	2.5%	-7.34 [-12.62, -2.06]
JN4043	3.67	2.36	3	9.33	0.94	3	3.5%	-5.66 [-8.53, -2.79]
JN4060	8	0.82	3	9	0.82	3	4.0%	-1.00 [-2.31, 0.31]
JN4066	5.33	1.7	3	7.33	2.62	3	3.2%	-2.00 [-5.53, 1.53]
JN4072	7.33	2.05	3	23.67	1.89	3	3.4%	-16.34 [-19.50, -13.18]
JN4073	3	2	2	18.5	0.5	2	3.5%	-15.50 [-18.36, -12.64]
JN4078	16.67	2.05	3	16	2.83	3	3.1%	0.67 [-3.28, 4.62]
JN4080	1	1.41	3	0.67	0.47	3	3.9%	0.33 [-1.35, 2.01]
JN5001	5.33	0.94	3	3.33	1.25	3	3.9%	2.00 [0.23, 3.77]
JN5002	16.5	1.5	2	19	2	2	3.3%	-2.50 [-5.96, 0.96]
JN5003	19.33	1.25	3	19.67	0.47	3	4.0%	-0.34 [-1.85, 1.17]
JN5004	17.33	0.47	3	21	4.08	3	2.8%	-3.67 [-8.32, 0.98]
JN5005	0.33	0.47	3	1	0.82	3	4.1%	-0.67 [-1.74, 0.40]
JN5006	10	4.32	3	11	3.56	3	2.2%	-1.00 [-7.33, 5.33]
JN5010	24.33	2.05	3	24.67	1.7	3	3.5%	-0.34 [-3.35, 2.67]
JN5011	3.67	0.94	3	5	0.85	3	4.0%	-1.33 [-2.76, 0.10]
JN5012	24.67	0.47	3	26.67	0.47	3	4.1%	-2.00 [-2.75, -1.25]
JN5014	14.67	4.71	3	16	2.16	3	2.3%	-1.33 [-7.19, 4.53]
JN5015	9.67	2.49	3	8.67	0.47	3	3.5%	1.00 [-1.87, 3.87]
JN5018	10.33	0.47	3	13	2.16	3	3.7%	-2.67 [-5.17, -0.17]
JN5026	3.61	2.94	3	6.08	4.97	3	2.1%	-2.47 [-9.00, 4.06]
JN5027	3.67	0.47	3	1	0.82	3	4.1%	2.67 [1.60, 3.74]
Total (95% CI)			**85**			**88**	**100.0%**	**-2.98 [-4.32, -1.65]**

Heterogeneity: Tau² = 11.08; Chi² = 352.81, df = 29 (P < 0.00001); I² = 92%
Test for overall effect: Z = 4.38 (P < 0.0001)

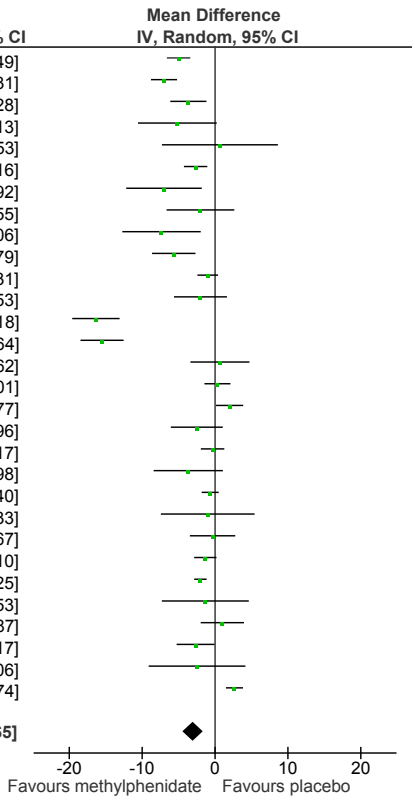


Figure 4.10.4.7 Teacher ratings of hyperactivity on the Conners scale

Study or Subgroup	Methylphenidate Mean	SD	Total	Placebo Mean	SD	Total	Weight	Mean Difference IV, Random, 95% CI
JN4001	0.33	0.47	3	4	1.41	3	14.8%	-3.67 [-5.35, -1.99]
JN4004	9	3.56	3	10.33	1.89	3	3.4%	-1.33 [-5.89, 3.23]
JN4008	6	1.63	3	7.33	1.25	3	10.0%	-1.33 [-3.65, 0.99]
JN4018	6.33	0.47	3	7.67	2.05	3	9.7%	-1.34 [-3.72, 1.04]
JN4019	11	2.94	3	12	2.97	3	3.2%	-1.00 [-5.73, 3.73]
JN4023	1.33	1.25	3	1.67	1.25	3	12.1%	-0.34 [-2.34, 1.66]
JN4027	5.33	1.25	3	6	1.63	3	10.0%	-0.67 [-2.99, 1.65]
JN4033	3.67	1.25	3	5.33	3.3	3	4.3%	-1.66 [-5.65, 2.33]
JN4034	8.33	2.62	3	9.67	1.89	3	5.0%	-1.34 [-5.00, 2.32]
JN4043	7	1.63	3	7.33	2.62	3	5.4%	-0.33 [-3.82, 3.16]
JN4049	10.33	3.25	3	13.67	1.7	3	4.0%	-3.34 [-7.49, 0.81]
JN4054	2.67	1.09	3	2.33	0.47	3	18.2%	0.34 [-1.00, 1.68]
Total (95% CI)			**36**			**36**	**100.0%**	**-1.22 [-2.11, -0.33]**

Heterogeneity: Tau² = 0.67; Chi² = 15.55, df = 11 (P = 0.16); I² = 29%
Test for overall effect: Z = 2.68 (P = 0.007)

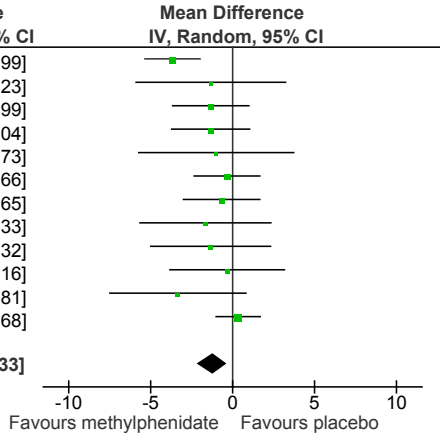


Figure 4.10.4.8 Parent ratings of hyperactivity on the Conners scale

Study or Subgroup	Methylphenidate Mean	SD	Total	Placebo Mean	SD	Total	Weight	Mean Difference IV, Random, 95% CI
JN4001	8.5	3.5	2	10.5	1.5	2	5.0%	-2.00 [-7.28, 3.28]
JN4019	8.44	0.94	3	11.67	3.3	3	6.7%	-3.23 [-7.11, 0.65]
JN4023	4.33	2.05	3	5	1.41	3	8.2%	-0.67 [-3.49, 2.15]
JN4043	14.67	0.47	3	17.67	0.94	3	10.4%	-3.00 [-4.19, -1.81]
JN4046	2.5	1.5	2	1.5	1.5	2	8.0%	1.00 [-1.94, 3.94]
JN4054	1	1	2	1.67	1.25	3	9.4%	-0.67 [-2.65, 1.31]
JN5004	5.33	0.47	3	5.67	0.94	3	10.4%	-0.34 [-1.53, 0.85]
JN5005	3.33	0.47	3	0.33	0.47	3	10.8%	3.00 [2.25, 3.75]
JN5006	2.33	1.7	3	8	2.83	3	6.9%	-5.67 [-9.41, -1.93]
JN5015	13	2.83	3	11.67	1.7	3	6.9%	1.33 [-2.41, 5.07]
JN5018	9.33	2.05	3	9	2.45	3	7.0%	0.33 [-3.28, 3.94]
JN5027	3.33	0.47	3	3.33	0.94	3	10.4%	0.00 [-1.19, 1.19]
Total (95% CI)			**33**			**34**	**100.0%**	**-0.65 [-2.24, 0.94]**

Heterogeneity: Tau² = 5.94; Chi² = 95.50, df = 11 (P < 0.00001); I² = 88%
Test for overall effect: Z = 0.80 (P = 0.42)

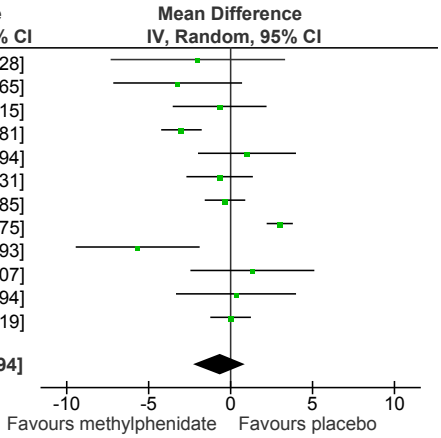


Figure 4.10.4.9 Teacher ratings of ADHD index on the Conners scale

Study or Subgroup	Methylphenidate Mean	SD	Total	Placebo Mean	SD	Total	Weight	Mean Difference IV, Random, 95% CI
JN4001	5	0.82	3	16	5.1	3	7.5%	-11.00 [-16.85, -5.15]
JN4004	20.67	3.3	3	26.67	3.86	3	7.7%	-6.00 [-11.75, -0.25]
JN4008	13	2.94	3	16.33	2.36	3	10.1%	-3.33 [-7.60, 0.94]
JN4018	16.67	3.4	3	21.67	1.25	3	10.5%	-5.00 [-9.10, -0.90]
JN4019	25	3	2	24.67	2.05	3	9.2%	0.33 [-4.43, 5.09]
JN4023	8.33	1.7	3	5.33	3.3	3	10.3%	3.00 [-1.20, 7.20]
JN4027	14	2.94	3	13.33	4.92	3	6.7%	0.67 [-5.82, 7.16]
JN4033	14	2.94	3	19.33	4.03	3	7.8%	-5.33 [-10.97, 0.31]
JN4034	21.67	3.77	3	27	2.16	3	9.0%	-5.33 [-10.25, -0.41]
JN4043	17	4.97	3	18.67	2.87	3	6.7%	-1.67 [-8.16, 4.82]
JN4049	23	5.34	2	25.67	3.59	3	4.7%	-2.67 [-11.11, 5.77]
JN4054	14	3.27	3	15	2.16	3	9.8%	-1.00 [-5.43, 3.43]
Total (95% CI)			**34**			**36**	**100.0%**	**-3.00 [-5.16, -0.84]**

Heterogeneity: Tau² = 7.22; Chi² = 22.54, df = 11 (P = 0.02); I² = 51%
Test for overall effect: Z = 2.72 (P = 0.007)

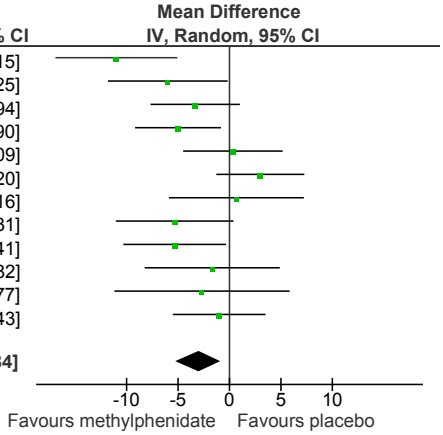


Figure 4.10.4.10 Parent ratings of ADHD index on the Conners scale

Study or Subgroup	Methylphenidate Mean	SD	Total	Placebo Mean	SD	Total	Weight	Mean Difference IV, Random, 95% CI
JN4001	5	0.82	3	16	1.16	3	11.2%	-11.00 [-12.61, -9.39]
JN4004	20.67	3.3	3	26.67	3.86	3	8.9%	-6.00 [-11.75, -0.25]
JN4008	13	2.94	3	16.33	2.36	3	9.9%	-3.33 [-7.60, 0.94]
JN4018	16.67	3.4	3	21.67	1.25	3	10.0%	-5.00 [-9.10, -0.90]
JN4019	28	3	2	24.67	2.05	3	9.6%	3.33 [-1.43, 8.09]
JN4023	8.33	1.7	3	5.33	3.3	3	9.9%	3.00 [-1.20, 7.20]
JN4027	14	2.94	3	13.33	0.33	3	10.5%	0.67 [-2.68, 4.02]
JN4033	14	2.94	3	19.33	0.67	3	10.4%	-5.33 [-8.74, -1.92]
JN4034	21.67	3.77	3	27	2.16	3	9.5%	-5.33 [-10.25, -0.41]
JN4043	17	2.12	3	18.67	2.87	3	10.1%	-1.67 [-5.71, 2.37]
Total (95% CI)			**29**			**30**	**100.0%**	**-3.14 [-6.78, 0.50]**

Heterogeneity: Tau² = 30.14; Chi² = 91.09, df = 9 (P < 0.00001); I² = 90%
Test for overall effect: Z = 1.69 (P = 0.09)

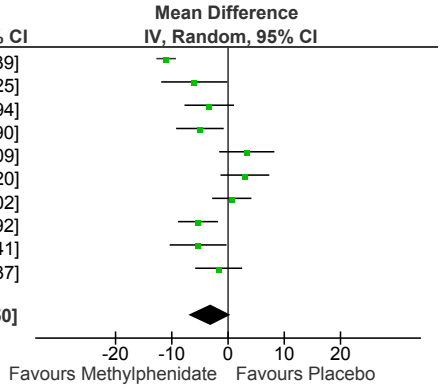


Figure 4.10.4.11 Teacher reports of total scores on ADHD rating scales

Study or Subgroup	Methylphenidate Mean	SD	Total	Placebo Mean	SD	Total	Weight	Mean Difference IV, Random, 95% CI
Payton1	5.5	2.12	2	22.5	4.95	2	31.7%	-17.00 [-24.46, -9.54]
Zwaigenbaum1	19.7	4.98	5	22.2	4.54	5	33.7%	-2.50 [-8.41, 3.41]
Zwaigenbaum2	13.33	2.65	3	10.5	3.62	3	34.6%	2.83 [-2.25, 7.91]
Total (95% CI)			**10**			**10**	**100.0%**	**-5.26 [-15.96, 5.45]**

Heterogeneity: Tau² = 79.47; Chi² = 18.61, df = 2 (P < 0.0001); I² = 89%
Test for overall effect: Z = 0.96 (P = 0.34)

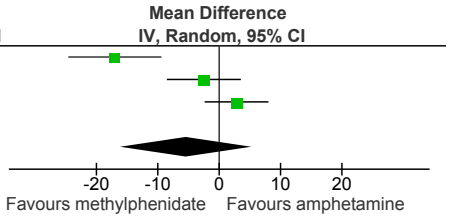
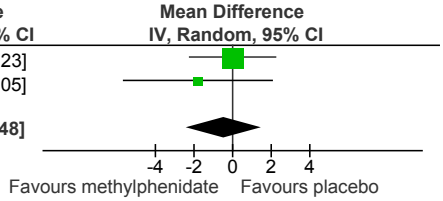


Figure 4.10.4.12 Parent reports of total scores on ADHD rating scales

Study or Subgroup	Methylphenidate Mean	SD	Total	Placebo Mean	SD	Total	Weight	Mean Difference IV, Random, 95% CI
Zwaigenbaum1	16.1	1.75	5	16.1	1.84	5	74.9%	0.00 [-2.23, 2.23]
Zwaigenbaum2	10	3.2	3	11.8	1.15	3	25.1%	-1.80 [-5.65, 2.05]
Total (95% CI)			**8**			**8**	**100.0%**	**-0.45 [-2.38, 1.48]**

Heterogeneity: Tau² = 0.00; Chi² = 0.63, df = 1 (P = 0.43); I² = 0%
Test for overall effect: Z = 0.46 (P = 0.65)



4.11APPENDICES

4.11.1 MEDLINE Search Strategy (amphetamine)

1. exp amphetamines/
2. exp central nervous system stimulants/
3. (amphetamine* or dexamphetamine* or methamphetamine* or dextroamphetamine* or lisdexamphetamine*).mp.
4. or/1-3
5. Attention Deficit Disorder with Hyperactivity/
6. exp hyperkinesis/
7. exp child behavior disorders/
8. ((attention adj deficit adj disorder) or hyperactive* or inattentive* or hyperkin*).mp.
9. (adhd or addh or adhs).mp.
10. (minimal* brain adj (dysfunct* or disorder* or damage*)).tw.
11. or/5-10
12. (child* or adolescent or infan*).mp.
13. 4 and 11 and 12
14. N-of-1.tw.
15. (individual* adj2 trial*).tw.
16. IMET*.tw.
17. Reversal design*.tw.
18. Alternating treatment design*.tw.
19. Multiple schedule design*.tw.
20. (ATD or ABA or ABAB*).tw.
21. (Multi-crossover or multiple crossover).tw.
22. Single system design*.tw.
23. (single adj (subject or patient or case) adj3 (trial* or design)).tw.
24. individuali#ed medication effectiveness test*.tw.
25. patient* as their own control*.tw.
26. or/14-25
27. 13 and 26

4.11.2 MEDLINE Search Strategy (Methylphenidate)

1. exp methylphenidate/
2. exp central nervous system stimulants/
3. (methylphenidate* or ritalin* or mph* or Concerta or equasym or methylin).mp.
4. or/1-3
5. Attention Deficit Disorder with Hyperactivity/
6. exp hyperkinesis/
7. exp child behavior disorders/
8. ((attention adj deficit adj disorder) or hyperactive* or inattentive* or hyperkin*).mp.
9. (adhd or addh or adhs).mp.
10. (minimal* brain adj (dysfunct* or disorder* or damage*)).tw.
11. or/5-10
12. (child* or adolescent or infan*).mp.
13. 4 and 11 and 12
14. N-of-1.tw.

15. (individual* adj2 trial*).tw.
16. IMET*.tw.
17. Reversal design*.tw.
18. Alternating treatment design*.tw.
19. Multiple schedule design*.tw.
20. (ATD or ABA or ABAB*).tw.
21. (Multi-crossover or multiple crossover).tw.
22. Single system design*.tw.
23. (single adj (subject or patient or case) adj3 (trial* or design)).tw.
24. individuali#ed medication effectiveness test*.tw.
25. patient* as their own control*.tw.
26. or/14-25
27. 13 and 26

Chapter 5: Amphetamine and methylphenidate for pediatric ADHD: A combined meta-analysis of N-of-1 trial data with RCT data

Salima Punja, Christopher H Schmid, Lisa Hartling, Liana Urichuk, Catherine J Nikles, Sunita Vohra

Submitted to: Journal of Clinical Epidemiology

5.1 ABSTRACT

Objectives: To assess how the inclusion of N-of-1 trial data into randomized controlled trial (RCT) meta-analyses impacts the magnitude and precision of yielded treatment effects, using amphetamines and methylphenidate for pediatric ADHD as a model.

Study Design and Setting: We combined the N-of-1 and RCT data generated from four previously conducted systematic reviews using parent and teacher ratings of hyperactivity/impulsivity as the outcome. Data was combined as standardized mean differences assuming a random effects model. The amphetamine and methylphenidate evidence were synthesized separately.

Results: We found that the inclusion of N-of-1 trial data in the meta-analysis impacted both magnitude and precision. The addition of the N-of-1 trial data narrowed the confidence intervals in all four comparisons as compared to the treatment effect yielded by RCT-only data. Furthermore, the addition of N-of-1 trials changed the overall treatment effects yielded by the RCT-only meta-analyses from statistically non-significant to statistically significant in one of the four comparisons.

Conclusions: If the overall goal of a meta-analysis is to synthesize all available evidence on a given topic, then N-of-1 trials should be included. This study shows that aggregate N-of-1 trials are comparable with RCT data and that it is possible to combine N-of-1 trial data with RCT data as well as the potential merits of this approach. Furthermore, when meta-analyzed, N-of-1 trials can produce population treatment effects comparable to those yielded by RCTs.

5.2. BACKGROUND

N-of-1 trials are multiple crossover trials conducted in single individuals and are often randomized and blinded (1). They allow for rigorous scientific investigation of the effectiveness of a particular treatment for an individual patient and promote evidence-based medicine (i.e. the integration of the best available evidence with clinical expertise and patient values). N-of-1 trials are applicable to stable, chronic conditions, where the treatment in question has a rapid onset and offset of action (1). They have been used to assess treatments in a number of conditions including arthritis, fibromyalgia and attention deficit hyperactivity disorder (ADHD).

N-of-1 trials offer particular advantages: i) unlike standard crossover trials where participants typically receive both the intervention and placebo only once, N-of-1 trial participants receive both intervention and placebo multiple times. This increases the power of the study for each individual participant; ii) N-of-1 trials offer direct evidence about treatment benefit to a patient, rather than the population mean yielded by randomized controlled trials (RCTs) which may or may not be applicable to a specific individual; iii) N-of-1 trials are methodologically rigorous; iv) N-of-1 trials promote improved patient safety by limiting therapies to only those that are demonstrated effective for a particular individual (i.e. reduction in polypharmacy); and v) when a series of N-of-1 trials of the same intervention is conducted in similar patients, with identical outcome measures, the results may be pooled for meta-analyses and compared with estimates of efficacy from an RCT. Despite the many advantages offered by N-of-1 trials, they are often excluded from systematic reviews and meta-analyses.

With the objective of identifying, appraising, and including evidence from all participants in an RCT, N-of-1 trials, which are a subset of RCTs, may serve as an additional source of data to be included in systematic reviews and meta-analyses. Previously conducted systematic reviews and meta-analyses synthesized a series of N-of-1 trials that assessed amphetamine and methylphenidate for pediatric ADHD (2, Chapter 4). These syntheses revealed the ability of N-of-1 trials to provide estimates of treatment effect both at the individual and aggregate patient level. In this study, we are particularly interested in assessing the impact of their inclusion into meta-analyses that have only included RCT data, using pediatric ADHD as a model. We hypothesize their inclusion may impact both the magnitude and precision of estimated treatment effects yielded by the meta-analyses compared to those restricted to RCT data alone. To test this hypothesis, we combined RCT and N-of-1 data into a single meta-analysis for two different ADHD treatments, methylphenidate and amphetamine. To our knowledge this is the first meta-analysis of this kind, whereby RCT and N-of-1 data are combined.

5.3. OBJECTIVES

The primary outcome of this meta-analysis is the change in the ADHD symptom hyperactivity/impulsivity according to parent and teacher ratings.

5.4 METHODS

5.4.1 Data collection

i. Amphetamine data

The N-of-1 data used in the amphetamine meta-analysis come from a previously

conducted systematic review and meta-analysis of amphetamines for pediatric ADHD (2, Chapter 4). Each N-of-1 trial included at least 2 treatment pairs in which each pair consisted of amphetamine and placebo.

The RCT data used for the amphetamine meta-analysis comes from a previously conducted systematic review (3, Chapter 3).

ii. Methylphenidate data

The N-of-1 data used in the methylphenidate meta-analysis come from a previously conducted systematic review and meta-analysis of methylphenidate for pediatric ADHD (2, Chapter 4). Each N-of-1 trial included at least 2 treatment pairs in which each pair consisted of methylphenidate and placebo.

The RCT data used for the methylphenidate meta-analysis comes from a previously conducted systematic review (4). Since this review did not report which specific studies were included in their meta-analyses, we contacted the author to retrieve this information, but it was no longer available. We therefore retrieved their included studies, re-extracted on our outcome of interest and re-ran the meta-analysis in accordance with the methods described in Chapter 4 of this dissertation.

5.4.2 Analysis

The N-of-1 and RCT data have already been meta-analyzed separately (see [2, Chapter 4] for a full description of methods), which produced mean differences for each intervention. In this meta-analysis we combined the N-of-1 and RCT evidence using standardized mean differences (SMDs) assuming a random-effects model. The amphetamine and methylphenidate evidence were synthesized separately.

5.5 RESULTS

5.5.1 Amphetamine

Teacher ratings

We used data from 26 N-of-1 trials (i.e. 26 participants), which provided 71 observations in the amphetamine arm and 71 observations in the placebo arm (since each participant provides 2-3 data points/intervention group). We combined this N-of-1 data with data from 1 RCT (n=70) which yielded a statistically significant SMD of -0.69 (95% CI -1.01 to -0.37) in favor of amphetamine (Figure 5.8.1). The addition of N-of-1 trials diminished the magnitude of the treatment effect yielded by RCT data alone by 0.44 standard deviations (SDs) and narrowed the confidence interval by 0.37 units.

Parent ratings

We combined data from 30 N-of-1 trials, which provided 88 observations in the amphetamine arm and eighty-eight observations in the placebo arm with data from 3 RCTs (n=417) and found a statistically significant SMD of -0.57 (95% CI -0.75 to -0.34) in favor of amphetamine (Figure 5.8.2). Both the aggregate n-of-1 trial data and aggregate RCT data yielded similar results. The addition of N-of-1 trials diminished the magnitude of treatment effect yielded by the RCT-only data by 0.06 SDs, and narrowed the confidence interval by 0.07 units.

5.5.2 Methylphenidate

Teacher ratings

We used data from thirty N-of-1 trials, which provided 82 observations in the methylphenidate arm and 83 observations in the placebo arm and combined this with data from 2 RCTs (n=126) and observed a statistically significant SMD -0.40 (95% CI -0.66

to -0.14) in favor of methylphenidate (Figure 5.8.3). Although the inclusion of N-of-1 trials did not impact the magnitude of the overall treatment effect yielded by the RCT data alone, they did affect the precision by narrowing the confidence interval by a factor of 0.24 units.

Parent ratings

We combined data from thirty N-of-1 trials, which provided 85 observations in the methylphenidate arm and 88 observations in the placebo arm with data from four RCTs (n=485) and found a statistically significant SMD of -0.54 (95% CI -0.86 to -0.22) in favor of methylphenidate (Figure 5.8.4). The impact of the N-of-1 trials on the magnitude was negligible; however, their inclusion increased the precision of the overall treatment effect by a factor of 0.48 units and changed the insignificant treatment effect yielded by RCT-only meta-analysis to a statistically significant one in favour of methylphenidate.

5.6 DISCUSSION

The primary purpose of this meta-analysis was to assess the impact of N-of-1 trial data on aggregated data in terms of magnitude and precision by combining the data into a single meta-analysis.

We found that the inclusion of N-of-1 trial data in the meta-analysis impacted both magnitude and precision. The addition of the N-of-1 trial data narrowed the confidence intervals in all four comparisons as compared to the treatment effect yielded by RCT-only data. Furthermore, the addition of N-of-1 trials changed the overall treatment effects yielded by the RCT-only meta-analyses from statistically non-significant to statistically significant in one of the outcomes (Figure 5.8.3). Of particular importance, we are able to

assess the robustness of the conclusions drawn from N-of-1-only analyses by comparing it with the RCT-only analyses. Our results show that the N-of-1 data were congruent with the results of the aggregate data in terms of magnitude and precision of treatment effect, which is an important step to validating the potential of N-of-1 trials to contribute to population-based treatment effects that are comparable to RCTs. Our data suggests that in instances where sufficient or rigorous RCT data are unavailable, one may be able to draw valid conclusions based solely on the results of aggregate N-of-1 data.

Although RCTs are considered the gold standard in assessing treatment effect, as they have high internal validity (i.e. the reduction or elimination in possible sources of bias), they have been criticized for their external validity (i.e. lack of generalizability) (5, 6). Lack of external validity/relevance is a shared criticism of both RCTs and systematic reviews (6, 7). While homogeneous populations maximize the ability to detect treatment effect, they limit the applicability of study findings to the diverse patient populations seen clinically. For example, rigid eligibility criteria may limit RCT enrolment to less than 10% of individuals with the disease in question (6). Unlike RCTs, N-of-1 trials can be tailored to individual patients, resulting in fewer exclusion criteria. As an example, while participants included in both the RCTs and N-of-1 trials were pediatric and had a clinical diagnosis of ADHD, the N-of-1 trials included in this paper reported no exclusion criteria, while the included amphetamine and methylphenidate RCTs report a total of eight and fifteen exclusion criteria respectively. The N-of-1 trials were thus able to promote external validity without compromising internal validity. As a result, the meta-analysis of N-of-1 trials may actually yield a more accurate treatment effect estimate that

represents what would happen in the 'real world' and potentially provides clinicians with relevant guidance regarding what treatment effect to expect in a patient who does not meet RCT eligibility criteria.

N-of-1 trials remain underutilized and more research is required on how to appropriately combine a series of N-of-1 trials with population-based data. The methodology used here was adopted from what is typically done when combining individual patient data (IPD) with aggregate data. The two most common methods include a one-step and two-step approach. The more popular two-step approach was utilized in this review and involves reducing the IPD (in this case the series of N-of-1 trials) to aggregate data and then conducting a standard meta-analysis of the aggregate data (8). In contrast, the one-step approach is much more complex and involves the use of multilevel modeling. Although the one-step method has the potential to consider study- and patient-level covariates in the effect estimate, it is quite complex and has had little statistical assessment, therefore, further research is needed to validate this approach (8).

A limitation of our analysis is that we were unable to obtain all of the primary N-of-1 data. Just as with publication bias in RCTs, a synthesis based on only N-of-1 trials may be biased if the unavailability of individual patient data is related to study results (i.e. data are not missing at random). However, by supplementing the N-of-1 data with RCT data (and vice versa), we have included as many participants as possible in this combined meta-analysis reducing the potential for bias in the overall effect estimate. We believe this is consistent with the philosophy of systematic reviews to be comprehensive and

utilize all available data.

Although the clinical focus of this paper is the treatment of ADHD, the implications of the results may be extended to a range of interventions and conditions that offer a large pool of N-of-1 trial data available for synthesis. While the primary value of N-of-1 trials lay in their ability to generate treatment effect estimates for individual patients, we must not overlook any secondary benefits that can be derived from N-of-1 trials (i.e. combining them to produce population-based estimates). Moreover, since N-of-1 trials represent a subset of RCTs, they should contribute to the overall picture of treatment effect in a systematic review and meta-analysis. By utilizing both N-of-1 and RCT data in this meta-analysis we are: i) promoting a comprehensive approach to include all available data; ii) achieving more precise effect estimates; and iii) increasing the overall power of the meta-analysis, thereby decreasing the likelihood of drawing false conclusions. Furthermore, as the movement towards IPD becomes the gold-standard for systematic reviews and meta-analyses (9), N-of-1 trials will play an important role as a type of IPD. While single N-of-1 trials provide information regarding treatment effectiveness in single individuals, the congruency of effect estimates yielded by the RCT-only and a series of N-of-1 trial-only data in this study show that a series of N-of-1 trials also have the potential to provide meaningful group estimates of treatment effect. If the overall goal of a meta-analysis is to synthesize all available evidence on a given topic, then N-of-1 trials should be included. This study shows it is possible to combine N-of-1 trial data with RCT data as well as the potential merits of this approach. Further considerations regarding the most appropriate statistical methods of combining these data is required.

5.7 REFERENCES

1. Guyatt G, Keller J, Jaeschke R, Rosenbloom D, Adachi JD, Newhouse MT. The N-of-1
   randomized controlled trial: Clinical usefulness. Our three-year experience. *Annals of Internal Medicine* 1990; 112:293–9

2. Punja S, Xu D, Schmid CH, Hartling L, Urichuk L, Nikles CJ, Vohra S. Amphetamines and methylphenidate for pediatric ADHD: A systematic review and meta-analysis of N-of-1 evidence. *Journal of Clinical Epidemiology* [in press].

3. Punja S, Shamseer L, Hartling L, Urichuk L, Vandermeer B, Nikles CJ, Vohra S. Amphetamines for attention deficit/hyperactivity disorder in children and adolescents. *Cochrane Database of Systematic Reviews* [in press].

4. Schachter HM, Pham B, King J, Langford S, Moher D. How efficacious and safe is short-acting methylphenidate for the treatment of attention-deficit hyperactivity disorder in children and adolescents? A meta-analysis. *Canadian Medical Association Journal* 2001;165(11):1475-1488.

5. Fortin M, Dionne J, Pinho G, Gignac J, Almirall J, Lapointe L. Randomized controlled trials: Do they have external validity for patients with multiple comorbidities? *Annals of Family Medicine* 2006; 4(2):101-108.

6. Rothwell PM. External validity of randomized controlled trials: “to whom do the results of this trial apply?” *Lancet* 2005; 365(9453):82-93.

7. Feinstein AR, Horwitz RI. Problems in the “evidence” of ‘evidence-based medicine”. *American Journal of Medicine* 1997;103:529-535.

8. Riley RD, Lambert PC, Staessen JA, et al. Meta-analysis of continuous outcomes combining individual patient data and aggregate data. *Statistics in Medicine* 2008;27:1870-93.

9. Stewart LA, Parmar MK. Meta-analysis of the literature or of individual patient data: is there a difference? *The Lancet* 1993; 341:418–422.

5.8 FIGURES

Figure 5.8.1 Amphetamine for pediatric ADHD: Teacher ratings

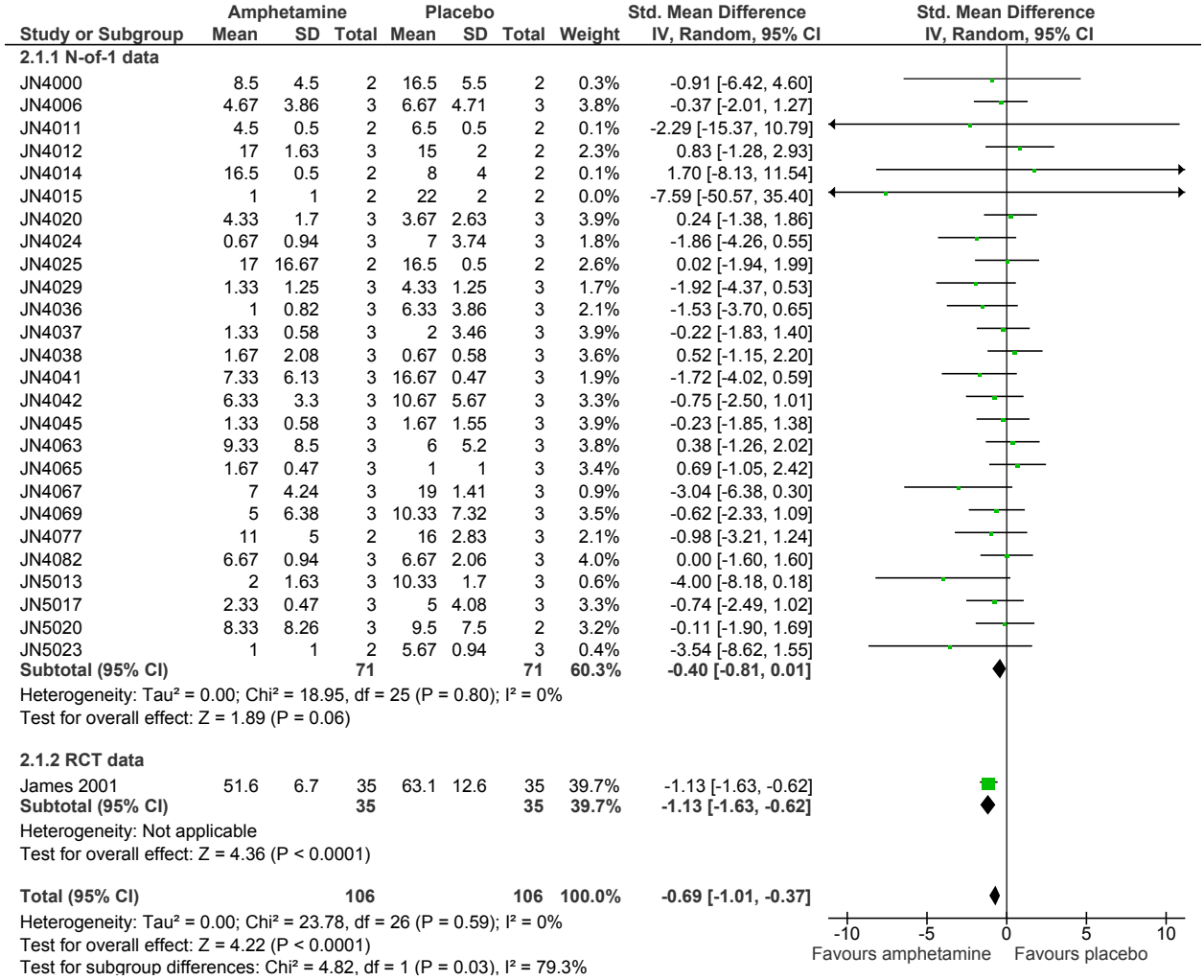


Figure 5.8.2 Amphetamine for pediatric ADHD: Parent ratings

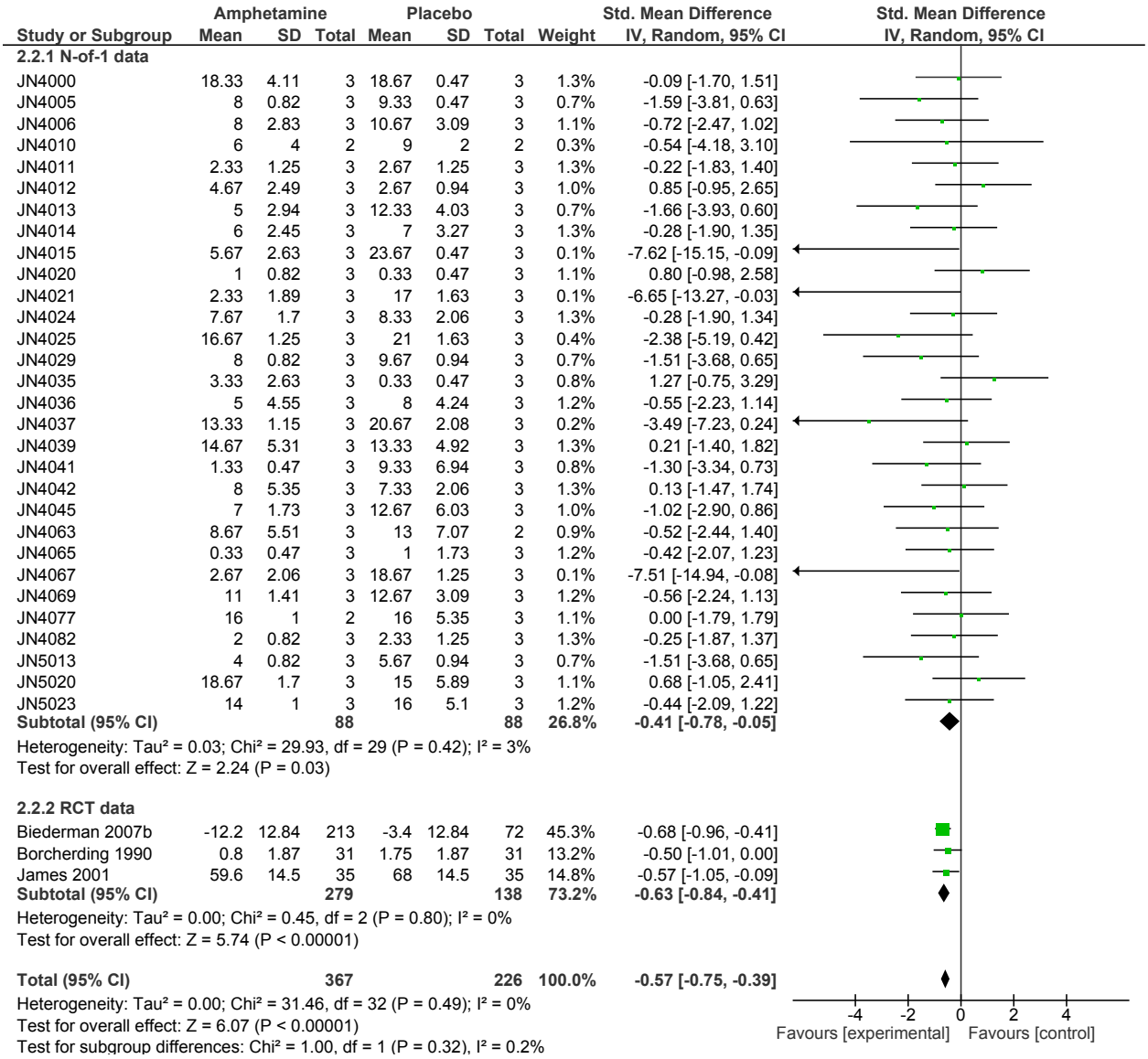


Figure 5.8.3 Methylphenidate for pediatric ADHD: Teacher ratings

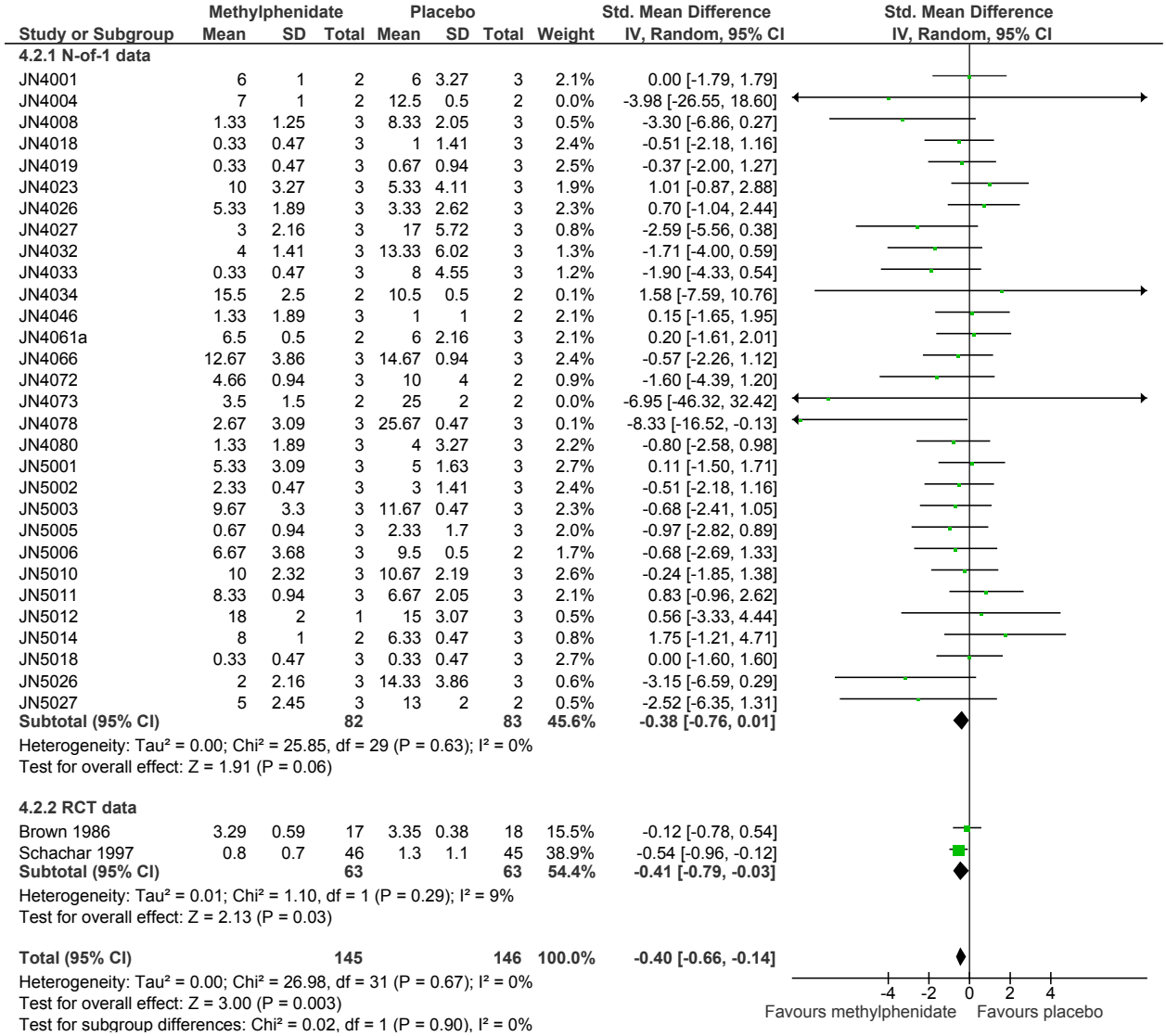
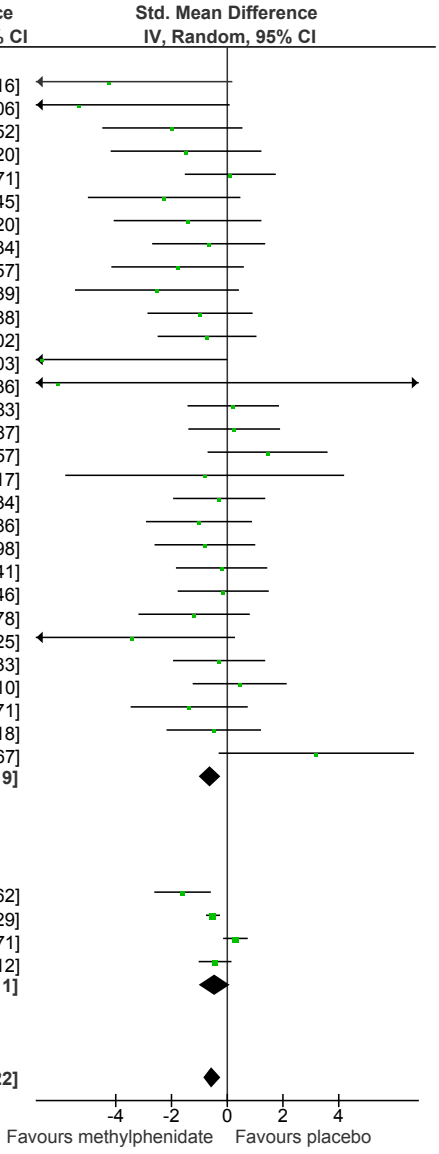


Figure 5.8.4 Methylphenidate for pediatric ADHD: Parent ratings

Study or Subgroup	Methylphenidate Mean	SD	Total	Placebo Mean	SD	Total	Weight	Std. Mean Difference IV, Random, 95% CI
4.4.1 N-of-1 data								
JN4001	1.67	0.47	3	6.67	1.25	3	0.5%	-4.24 [-8.63, 0.16]
JN4004	14	0.48	3	21	1.41	3	0.3%	-5.32 [-10.69, 0.06]
JN4008	7.67	0.47	3	11.33	2.05	3	1.5%	-1.97 [-4.45, 0.52]
JN4018	7.5	3.5	2	12.67	1.89	3	1.3%	-1.48 [-4.16, 1.20]
JN4019	18	2	3	17.33	6.65	3	3.2%	0.11 [-1.49, 1.71]
JN4023	2.67	0.94	3	5.33	0.94	3	1.3%	-2.26 [-4.97, 0.45]
JN4027	5	1	2	12	4.32	3	1.4%	-1.42 [-4.05, 1.20]
JN4033	10	3	2	12	1.63	3	2.2%	-0.67 [-2.67, 1.34]
JN4034	11.33	2.05	3	18.67	4.19	3	1.7%	-1.78 [-4.13, 0.57]
JN4043	3.67	2.36	3	9.33	0.94	3	1.1%	-2.52 [-5.43, 0.39]
JN4060	8	0.82	3	9	0.82	3	2.5%	-0.98 [-2.83, 0.88]
JN4066	5.33	1.7	3	7.33	2.62	3	2.7%	-0.72 [-2.47, 1.02]
JN4072	7.33	2.05	3	23.67	1.89	3	0.2%	-6.63 [-13.23, -0.03]
JN4073	3	2	2	18.5	0.5	2	0.0%	-6.08 [-40.51, 28.36]
JN4078	16.67	2.05	3	16	2.83	3	3.1%	0.22 [-1.40, 1.83]
JN4080	1	1.41	3	0.67	0.47	3	3.1%	0.25 [-1.37, 1.87]
JN5001	5.33	0.94	3	3.33	1.25	3	2.0%	1.45 [-0.68, 3.57]
JN5002	16.5	1.5	2	19	2	2	0.4%	-0.81 [-5.78, 4.17]
JN5003	19.33	1.25	3	19.67	0.47	3	3.1%	-0.29 [-1.91, 1.34]
JN5004	17.33	0.47	3	21	4.08	3	2.4%	-1.01 [-2.89, 0.86]
JN5005	0.33	0.47	3	1	0.82	3	2.7%	-0.80 [-2.58, 0.98]
JN5006	10	4.32	3	11	3.56	3	3.1%	-0.20 [-1.81, 1.41]
JN5010	24.33	2.05	3	24.67	1.7	3	3.1%	-0.14 [-1.75, 1.46]
JN5011	3.67	0.94	3	5	0.85	3	2.3%	-1.19 [-3.16, 0.78]
JN5012	24.67	0.47	3	26.67	0.47	3	0.7%	-3.40 [-7.06, 0.25]
JN5014	14.67	4.71	3	16	2.16	3	3.1%	-0.29 [-1.92, 1.33]
JN5015	9.67	2.49	3	8.67	0.47	3	3.0%	0.45 [-1.21, 2.10]
JN5018	10.33	0.47	3	13	2.16	3	2.1%	-1.37 [-3.44, 0.71]
JN5026	3.61	2.94	3	6.08	4.97	3	3.0%	-0.48 [-2.15, 1.18]
JN5027	3.67	0.47	3	1	0.82	3	0.8%	3.20 [-0.28, 6.67]
Subtotal (95% CI)			**85**			**88**	**57.9%**	**-0.60 [-1.01, -0.19]**
Heterogeneity: Tau² = 0.11; Chi² = 31.80, df = 29 (P = 0.33); I² = 9%								
Test for overall effect: Z = 2.90 (P = 0.004)								
4.4.2 RCT data								
Brown 1988	4.22	1.39	11	8.33	3.2	11	6.2%	-1.60 [-2.59, -0.62]
Fischer 1991	4.95	3.02	161	6.6	3.4	161	13.9%	-0.51 [-0.73, -0.29]
Schachar 1997	1.2	1.1	46	0.9	0.9	45	11.9%	0.30 [-0.12, 0.71]
Stein 1996	4.85	3	25	6.3	3.5	25	10.2%	-0.44 [-1.00, 0.12]
Subtotal (95% CI)			**243**			**242**	**42.1%**	**-0.45 [-1.01, 0.11]**
Heterogeneity: Tau² = 0.25; Chi² = 17.57, df = 3 (P = 0.0005); I² = 83%								
Test for overall effect: Z = 1.56 (P = 0.12)								
Total (95% CI)			**328**			**330**	**100.0%**	**-0.54 [-0.86, -0.22]**
Heterogeneity: Tau² = 0.18; Chi² = 50.16, df = 33 (P = 0.03); I² = 34%								
Test for overall effect: Z = 3.27 (P = 0.001)								
Test for subgroup differences: Chi² = 0.20, df = 1 (P = 0.66), I² = 0%								



Chapter 6: Conclusion

6.1 SUMMARY OF KEY FINDINGS

This dissertation consists of four chapters. In Chapter 2, the methods of design, analysis and meta-analysis of published N-of-1 trials were systematically reviewed. This review revealed that N-of-1 trials have been conducted in over 50 conditions, suggesting they are amenable to evaluate a variety of health conditions. Contrary to assumptions that N-of-1 trials, by definition, only have a single participant, the majority have been published as a series. Furthermore, the results show that the majority of N-of-1 trials utilized elements that maintain methodological rigour such as randomization, blinding, and formal outcome assessments. Similar to randomized controlled trials (RCTs), the reporting of N-of-1 trials has been demonstrated to be inadequate and could benefit from a reporting guideline. N-of-1 trials offer a number of advantages including: i) offering an individualized assessment; ii) giving patients/participants the opportunity to experience all interventions being assessed; iii) ensuring the results are directly relevant and applicable to the patients/participants themselves. This review revealed that these advantages are being realized in the published literature, and as such, the utilization of N-of-1 trials expands each year.

In Chapter 3, we undertook a systematic review and meta-analysis of amphetamines for attention deficit/hyperactivity disorder (ADHD) in children and adolescents. This review found that although amphetamines are effective at reducing the core symptoms of ADHD in the short-term, they were also associated with a higher risk of adverse events including decreased appetite, insomnia, and abdominal pain compared to placebo. Subgroup analysis revealed no difference in efficacy between the long-acting and short-acting preparations. Although the results of the subgroup analysis should be interpreted with

caution, they suggest that despite costing up to fifteen times more, long-acting preparations are comparable to their less expensive shorter-acting counterparts. Further research is needed to investigate this, including whether longer-acting formulations achieve the promise of greater compliance (which is the main advantage described in their marketing). The Chapter 3 systematic review was based solely on RCT data, which was used to inform Chapter 5 of this dissertation.

In Chapter 4, we evaluated how data from N-of-1 trials may be used in systematic reviews and meta-analyses by examining the effects of amphetamine and methylphenidate for pediatric ADHD. To our knowledge, a systematic review and meta-analysis of this kind has not previously been conducted. Our findings indicated that both amphetamines and methylphenidate were superior to placebo on most outcomes including teacher ratings of inattention and hyperactivity/impulsivity, as well as parent ratings of inattention and hyperactivity/impulsivity. Furthermore, by meta-analyzing N-of-1 evidence, we were able to measure both individual estimates of treatment effect, as well population estimates of treatment effect.

Chapter 5 included a combined meta-analysis of N-of-1 and RCT data. The objective of this meta-analysis was to assess how the inclusion of N-of-1 trial data into RCT meta-analyses impacts the magnitude and precision of yielded treatment effects, using amphetamine and methylphenidate for pediatric ADHD as a model. To our knowledge, a meta-analysis of this kind has not previously been conducted. We found that the inclusion of N-of-1 trial data in the meta-analysis impacted both magnitude and precision. The addition of the N-of-1 trial data narrowed the confidence intervals across all outcomes for both interventions. Furthermore, the addition of N-of-1 trials changed

the overall treatment effects yielded by the RCT-only meta-analyses from statistically non-significant to statistically significant in parent ratings of hyperactivity/impulsivity in favour of methylphenidate compared to placebo. Moreover, this review showed that the results yielded by aggregate N-of-1 trials do produce comparable results to aggregate RCT data.

6.2 LIMITATIONS

Although meta-analysis of N-of-1 trials is possible and can be used to produce both individual treatment effects as well as population treatment effects, the following should be considered when interpreting the findings of this dissertation. Given that the N-of-1 trial is an emerging area, optimal methods of meta-analysis must be appropriately assessed. In Chapter 3, methods utilized for RCT meta-analyses (specifically what is recommended for IPD meta-analyses) were extrapolated to the meta-analysis of N-of-1 trials; however, other methods have been proposed, specifically Bayesian techniques. Different proposed methods should be compared in order to further explore the strengths and limitations of each. Furthermore, while it is clear that N-of-1 trials can and have been used to assess a number of conditions, it is important to note that this dissertation focused solely on the effects of amphetamines and methylphenidate for pediatric ADHD. In order to assert the generalizability of the results and the validity of our findings, meta-analyses utilizing similar methodology should be conducted in areas where N-of-1 trials have been popular, including osteoarthritis, chronic pain, and sleep problems.

6.3 IMPLICATIONS FOR CLINICAL PRACTICE

Given the lack of rigorous evidence on long-term effectiveness, comparative effectiveness, and additive effectiveness (i.e. whether a patient received additional benefit

from combination therapy) as well as the narrow eligibility criteria of RCTs which exclude the majority of patients seen in routine clinical practice, clinical treatment decision-making is often based on clinical experience, the lowest level of evidence-based medicine. As such, the potential for N-of-1 trials in clinical practice is immense, allowing for rigorous scientific investigation of the effectiveness of a particular treatment for an individual patient, and promoting evidence-based medicine (i.e., the integration of the best available evidence with clinical expertise and patient values). N-of-1 trials are particularly useful in clinical practice because: i) they are ideal at assessing long-term therapy in chronic conditions; ii) they are ideal for patients with comorbid conditions and those using concurrent therapies (as in the majority of patients seen in clinical practice); iii) they are able to establish comparative effectiveness and additive effectiveness; iv) they promote personalized medicine by avoiding a 'one size fits all' approach to delivery of health care; v) they may help reduce polypharmacy and thus promote patient safety by limiting therapies to those with demonstrated effectiveness; and vi) they allow clinicians to develop evidence that helps promote evidence-based decisions in their individual practices.

Despite the numerous advantages of N-of-1 trials, they remain an underutilized tool in clinical care. A number of barriers must be overcome to allow for their widespread adoption. One of the potential barriers is that physicians often lack an awareness and understanding of the design, analysis, interpretation and benefits of N-of-1 trials. As such, appropriate training and support should be provided to physicians. This can include the development of an N-of-1 curriculum which can be implemented into the basic medical education training, or offered as online tutorials as continuing education (1).

Expanding the appeal of N-of-1 trials can also occur at academic conferences or via targeted marketing to clinics that would benefit the greatest such as family practices, pediatric practices, and specialty clinics that treat patients with chronic conditions.

Another potential barrier of conducting N-of-1 trials in clinical practice is that physicians may lack the necessary resources (e.g. collaborating pharmacist, statistical expertise) to conduct and evaluate N-of-1 trials. Establishing an N-of-1 trial service that works to either provide these resources to physicians or conducts N-of-1 trials on a referral basis would be valuable. Investigators have established N-of-1 trial services, such as Dr. Gordon Guyatt at McMaster University, Dr. Sunita Vohra at the University of Alberta, as well as a national service established at the University of Queensland. Unfortunately, while these services were successful, they all came to a close once their funding was terminated; therefore, a self-sustaining N-of-1 clinical trial service is necessary to overcome these barriers. Such a service could provide key support functions to clinicians and offer a range of services from simply providing online tools to assist physicians in setting up their own N-of-1 trials to a more comprehensive consultation service which works to i) identify the clinical question; ii) select the appropriate outcome tools; iii) assist in the design of the trial (e.g. development of the randomization code; preparing the trial interventions); and iv) provide statistical support. Thus, a successful service would involve a range of expertise including researchers, clinicians, pharmacists and statisticians.

Although N-of-1 trials have typically been used to evaluate effectiveness of a therapy in an individual patient, this thesis confirms that the majority of published N-of-1 trials are being conducted as a series for the same condition-intervention pair and can be meta-

analyzed to produce population treatment effects. As such, N-of-1 trials may be used to create predictive models to assist physicians in more accurate prescribing. By creating a database of conducted N-of-1 trials both in clinical care and research, physicians can refer to and determine which prognostic factors match with the most successful treatment option, ultimately resulting in enhanced patient care.

6.4 IMPLICATIONS FOR RESEARCH

The potential for N-of-1 trials is significant and, therefore, research into the most optimal methods of analysis and meta-analysis is needed. Bayesian techniques are garnering interest particularly in the area of N-of-1 trials since these methods can be used to both analyze and meta-analyze N-of-1 trials. Bayesian statistics are concerned with the probability of a parameter given the observed data. Its strengths lay in its ability to maximize the use of available information from each participant as well as its utilization of prior information being incorporated into the statistical model so that each conducted N-of-1 trial can inform the next; thus eliminating the need for sample size calculations and allowing for more efficient use of resources. The pitfalls of the Bayesian method is that it is quite complex, and in order to maximize its strength all of the parameters within the model need to be pre-specified. However, it is often the case that little prior information is known about these parameters. Further exploration of this method, including a comparison of the method conducted in this dissertation and how they compare to population estimates yielded by aggregate data is needed.

Publication bias may be an even larger concern in systematic reviews and meta-analysis of N-of-1 trials than it is for RCTs. It is extremely unlikely that all N-of-1 trials being

conducted are being published, particularly those that are being carried out for clinical purposes. As such, syntheses of N-of-1 trials may overlook a large subset of N-of-1 trials, yielding potentially biased results. Although there has not been any work done with respect to publication bias in N-of-1 trials, given their novelty, some work has been conducted in this area with respect to individual participant data (IPD) meta-analyses. One study found that those IPD meta-analyses which excluded unpublished literature yielded exaggerated treatment effects (2). Given that N-of-1 trials are a type of IPD, these conclusions may be extrapolated to syntheses of N-of-1 trials. A potential solution to capturing this subset of N-of-1 trials would be to establish a registry, similar to what is done with RCTs, in which both clinicians and researchers can register their protocols and report their results. Although this may seem unnecessary to clinicians carrying out N-of-1 trials for clinical care purposes and not research purposes, it must be made clear to them that by registering their N-of-1 trial information, the benefits of that N-of-1 trial can go beyond simply the treatment of their individual patients and be utilized for future quality improvement endeavors.

The issue of publication bias goes beyond simply excluding unpublished literature when it comes to meta-analyses of IPD and therefore of N-of-1 trials. The issue of data availability bias may also arise, whereby IPD is unavailable for some studies and available for others (3). This type of bias may further exaggerate yielded treatment effects. One study found that 52% of identified IPD meta-analyses could not obtain all the IPD requested due to data being lost or destroyed as well as trial authors being unreachable, unwilling to collaborate or simply unable to share their data (3). The

implementation of data sharing initiatives, promotion of transparent reporting through the use of the CONSORT Extension for N-of-1 trials (CENT) (4), as well as the use to trial registries should work towards reducing data availability bias in meta-analyses of N-of-1 trials.

6.5 REFERENCES

1. Kaplan HC, Gabler NB and the DEcIDE Methods Center N-of-1 Guidance Panel (Eslick I, Duan N, Kravitz R, Larson E, Pace W, Schmid C, Sim I, Vohra S). User Engagement, Training, and Support for Conducting N-of-1 trials. Agency for Healthcare Research and Quality; AHRQ Publication No.13(14)-EHC122-EF. Rockville, MD. February 2014

2. Burdett S, Stewart LA, Tierney JF. Publication bias and meta-analyses: a practical example. *International Journal of Technology Assessments in Health Care* 2003;19:129-34.

3. Ahmed I, Sutton AJ, Riley RD. Assessment of publication bias, selection bias, and unavailable data in meta-analyses using individual participant data: a database survey. *British Medical Journal* 2012;344:d7762.

4. Shamseer L, Sampson M, Bukutu C, Barrowman N, Altman D, Moher D, Vohra S. P.05.50. CONSORT Extension for N-of-1 trials (CENT) Guidelines. *BMC Complementary and Alternative Medicine* 2012; 12(Suppl 1): P410.

All sources used

1. Agency for Health Care Policy and Research. Diagnosis of Attention-Deficit/Hyperactivity Disorder. Summary, Technical Review. http://www.ahrq.gov/clinic/epcsums/adhdsutr.htm (accessed 1 May 2013).
2. Ahmed I, Sutton AJ, Riley RD. Assessment of publication bias, selection bias, and unavailable data in meta-analyses using individual participant data: a database survey. *British Medical Journal* 2012;344:d7762.
3. American Academy of Pediatrics, Subcommittee on Attention-Deficit/Hyperactivity Disorder, Steering Committee on Quality Improvement and Management. ADHD: Clinical Practice Guideline for the Diagnosis, Evaluation, and Treatment of Attention-Deficit/Hyperactivity Disorder in Children and Adolescents. Pediatrics 2011;128:1007-22.
4. American Psychiatric Association. Diagnostic and Statistical Manual of Mental Disorders (DSM-III-R). 3rd edition. Washington, DC: American Psychiatric Association, 1987.
5. American Psychiatric Association. Diagnostic and Statistical Manual of Mental Disorders (DSM-IV-TR). 4th text revision edition. Washington, DC: American Psychiatric Association, 2000.
6. Anderson SW, Bechara A, Damasio H, Tranel D, Damasio AR. Impairment of social and moral behavior related to early damage in human prefrontal cortex. *Natural Neuroscience* 1999;2:1032-7.

7. Anderson EE, Clement PW, Oettinger Jr. L. Methylphenidate compared with behavioral self-control in attention deficit disorder: Preliminary report. *Developmental and Behavioral Pediatrics* 1981;2(4):137-140.
8. Arnsten AFT. Stimulants: therapeutic actions in ADHD. *Neuropsychopharmacology* 2006;31:2376-83.
9. Barkley RA, Connor DF, Kwasnik D. Challenges to determining adolescent medication response in an outpatient clinical setting: comparing Adderall and methylphenidate for ADHD. *Journal of Attention Disorders* 2000;4(2):102-13.
10. Biederman J, Boellner SW, Childress A, Lopez FA, Krishnan S, Zhang Y. Lisdexamfetamine dimesylate and mixed amphetamine salts extended-release in children with ADHD: a double-blind, placebo-controlled, crossover analog classroom study. *Biological Psychiatry* 2007;62(9):970-6.
11. Biederman J, Krishnan S, Zhang Y, McGough JJ, Findling RL. Efficacy and tolerability of lisdexamfetamine dimesylate (NRP-104) in children with attention-deficit/hyperactivity disorder: a phase III, multicenter, randomized, double-blind, forced-dose, parallel-group study. *Clinical Therapeutics* 2007;29(3):450-63.
12. Biederman J, Lopez FA, Boellner SW, Chandler MC. A randomized, double-blind, placebo-controlled, parallel-group study of SLI381 (Adderall XR) in children with attention-deficit/hyperactivity disorder. *Pediatrics* 2002;110(2):258-66.
13. Borcherding BG, Keysor CS, Rapoport JL, Elia J, Amass J. Motor/vocal tics and compulsive behaviors on stimulant drugs: is there a common vulnerability? *Psychiatry Research* 1990;33(1):83-94.

14. Buck ML. Amphetamines in the treatment of attention-deficit/hyperactivity disorder. *Pediatric Pharmacotherapy* 2002;8(3):1-3.

15. Burdett S, Stewart LA, Tierney JF. Publication bias and meta-analyses: a practical example. *International Journal of Technology Assessments in Health Care* 2003;19:129-34.

16. Castells X, Ramos-Quiroga JA, Bosch R, Nogueira M, Casas M. Amphetamines for Attention Deficit Hyperactivity Disorder (ADHD) in adults. *Cochrane Database of Systematic Reviews* 2011, Issue 6. Art. No.: CD007813. DOI: 10.1002/14651858.CD007813.pub2.

17. Chalmers I. The Cochrane Collaboration: preparing, maintaining and disseminating systematic reviews of the effects of health care. *Annals of the New York Academy of Science* 1993;703:156-165.

18. Charach A, Dashti B, Carson P, Booker L, Lim CG, Lillie E, et al. Attention deficit hyperactivity disorder: effectiveness of treatment in at-risk preschoolers; long-term effectiveness in all ages; and variability in prevalence, diagnosis, and treatment. Agency for Healthcare Research and Quality 2011.

19. Cohen J. Statistical Power Analysis in Behavioral Sciences (2nd edition. Hillsdale (NJ): Lawrence Erlbaum Associates Inc., 1988.

20. Cook DJ, Mulrow CD, Haynes RB. Systematic Reviews: Synthesis of Best Evidence for Clinical Decisions. *Annals of Internal Medicine* 1997;126(5): 376-380. doi:10.7326/0003-4819-126-5-199703010-00006

21. Donnelly M, Rapoport JL, Ismond DR. Fenfluramine treatment of childhood attention deficit disorder with hyperactivity: a preliminary report. *Psychopharmacology Bulletin* 1986;22(1):152-4.

22. Donnelly M, Rapoport JL, Potter WZ, Oliver J, Keysor CS, Murphy DL. Fenfluramine and dextroamphetamine treatment of childhood hyperactivity. Clinical and biochemical findings. *Archives of General Psychiatry* 1989;46(3):205-12.

23. Duggan CM, Mitchell G, Nikles CJ, et al. Managing ADHD in general practice. N-of-1 trials can help! *Australian Family Physician* 2000;29:1205–1209.

24. Edwards L, Salant V, Howard VF, Brougher J, McLaughlin TF. Effectiveness of self-management on attentional behavior and reading comprehension for children with attention deficit disorder. *Child and Family Behavior Therapy* 1995;17(2):1-17.

25. Egger M, Smith GD, Schneider M, Junker C, Lengeler C, Antes G. Language bias in randomized controlled trials published in English and German. *The Lancet* 1997;350:326-9.

26. Elia J, Borcherding BG, Potter WZ, Mefford IN, Rapoport JL, Keysor CS. Stimulant drug treatment of hyperactivity: biochemical correlates. *Clinical Pharmacology and Therapeutics* 1990;48(1):57-66.

27. Elia J, Borcherding BG, Rapoport JL, Keysor CS. Methylphenidate and dextroamphetamine treatments of hyperactivity: are there true nonresponders? *Psychiatry Research* 1991;36(2):141-55.

28. Elia J, Welsh PA, Gullotta CS, Rapoport JL. Classroom academic performance: improvement with both methylphenidate and dextroamphetamine in ADHD boys. *Journal of Child Psychology and Psychiatry* 1993;34(5):785-804.

29. Fanton J, Waslick B, Harvey E. The 49th Annual National Institute of Mental Health (NIMH) New Clinical Drug Evaluation Unit (NCDEU) meeting Hollywood, Florida, June29-July 2, 2009: posters most relevant to child and adolescent psychopharmacology. *Journal of Child and Adolescent Psychopharmacology* 2009;19(6):786.
30. Faraone SV, Biederman J, Mick E. The age-dependent decline of attention-deficit/hyperactivity disorder: a meta-analysis of follow-up studies. *Psychological Medicine* 2006;36:159–165.
31. Feinstein AR, Horwitz RI. Problems in the "evidence" of 'evidence-based medicine". *American Journal of Medicine* 1997;103:529-535.
32. Findling RL, Childress AC, Cutler AJ, Gasior M, Hamdani M, Ferreira-Cornwell MC, et al. Efficacy and safety of lisdexamfetamine dimesylate in adolescents with attention-deficit/hyperactivity disorder. *Journal of the American Academy of Child and Adolescent Psychiatry* 2011;50(4):395-405.
33. Fortin M, Dionne J, Pinho G, Gignac J, Almirall J, Lapointe L. Randomized controlled trials: do they have external validity for patients with multiple comorbidities? *The Annals of Family Medicine* 2006;4(2):104-108.
34. Gabler NB, Duan N, Vohra S, Kravitz RL. N-of-1 trials in the medical literature: a systematic review. *Medical Care* Aug 2011;49(8):761-768.
35. Giblin GM, Strobel AL. Effect of lisdexamphetamine dimesylate on sleep in children with ADHD. *Journal of Attention Disorders* 2011;15(6):491-8.
36. Gillberg C, Melander H, Knorring AL, Janols LO, Thernlund G, Hagglof B, et al. Long-term stimulant treatment of children with attention-deficit hyperactivity

disorder symptoms. A randomized, double-blind, placebo-controlled trial. *Archives of General Psychiatry* 1997;54(9):857-64.

37. Gitterman DP, Hay Jr WW. That Sinking Feeling, Again? The State of National Institutes of Health Pediatric Research Funding, Fiscal Year 1992-2010. *Pediatric Research* 2008;64(5):462.
38. Glos J. [Amphetamines in the treatment of hyperactive and instable children]. *Ceskoslovenska pediatrie* 1973;28(10):559-60.
39. Guyatt G, Hayward R, Richardson WS, et al. Moving from evidence to action. In: Guyatt G,Rennie D, eds. The user's guides to the medical literature: a manual for evidence-based clinical practice. Chicago, IL: AMA Publications, 2002 p. 275–90.
40. Guyatt G, Jaeschke R, McGinn T. N-of-1 randomized controlled trials. In: Guyatt G, Rennie D, Meade MO, Cook DJ, editors. Users' Guides to the Medical Literature: A Manual for Evidence-Based Clinical Practice. 2nd ed. New York, NY: McGraw-Hill; 2008.
41. Guyatt G, Keller J, Jaeschke R, Rosenbloom D, Adachi JD, Newhouse MT. The N-of-1

randomized controlled trial: Clinical usefulness. Our three-year experience. *Annals of Internal Medicine* 1990; 112:293–9.

42. Guyatt G, Sackett D, Adachi J et al. A clinician's guide for conducting randomized trials in

individual patients. *Canadian Medical Association Journal* 1988; 139:497–503.

43. Hartling L, Ospina M, Liang Y, Dryden DM, Hooton N, Krebs Seida J, et al. Risk of bias versus quality assessment of randomized controlled trials: cross sectional study. *British Medical Journal* 2009;339:1-6.

44. Helsel WJ, Hersen M, Lubetsky MJ, Fultz SA, Sisson L, Harlovic CH. Stimulant drug treatment of four multihandicapped children using randomized single-case design. *Journal of the Multihandicapped Person* 1989;2(2):139-153.

45. Higgins JPT, Green S (editors). Cochrane Handbook for Systematic Reviews of Interventions. Chichester: John Wiley & Sons, 2011.

46. Hupp SD, Reitman D, Northup J, O'Callaghan P, LeBlanc M. The effects of delayed rewards, tokens, and stimulant medication on sportsmanlike behavior with ADHD-diagnosed children. *Behavior Modification* 2002;26(2):148-162.

47. Itil TM, Simeon J. Proceedings: Computerized EEG in the prediction of outcome of drug treatment in hyperactive childhood behavior disorders. *Psychopharmacology Bulletin* 1974;10(4):36.

48. Jaeschke R, Adachi J, Guyatt G, et al. Clinical usefulness of amitriptyline in fibromyalgia: the results of 23 N-of-1 randomized controlled trials. *Journal of Rheumatology* 1991;18:447–451.

49. Jadad AR. Randomized controlled trials: a user's guide. London, England: BMJ Books, 1998.

50. James RS, Sharp WS, Bastain TM, Lee PP, Walter JM, Czarnolewski M, et al. Double-blind, placebo-controlled study of single-dose amphetamine formulations in ADHD. *Journal of the American Academy of Child and Adolescent Psychiatry* 2001;40(11):1268-76.

51. Johannessen T, Petersen H, Kristensen P, Fosstvedt D. The controlled single subject trial. *Scandinavian Journal of Primary Health Care* 1991; 9:17–21.

52. Kamien M. The use of an N-of-1 randomized clinical trial in resolving therapeutic doubt. *Australian Family Physician* 1998;27(Suppl 2):S103-S105.

53. Kaplan HC, Gabler NB and the DEcIDE Methods Center N-of-1 Guidance Panel (Eslick I, Duan N, Kravitz R, Larson E, Pace W, Schmid C, Sim I, Vohra S). User Engagement, Training, and Support for Conducting N-of-1 trials. Agency for Healthcare Research and Quality; AHRQ Publication No.13(14)-EHC122-EF. Rockville, MD. February 2014

54. Kazdin AE. Single-Case Research Designs: Methods for Clinical and Applied Settings. New

York: Oxford University Press, 1982.

55. Kutcher SP. Assessing and treating attention deficit disorder in adolescents. The clinical application of a single-case research design. *The British Journal of Psychiatry* 1986;149(6):710-715.

56. LaRue Jr. RH, Northup J, Baumeister AA, Hawkins MF, Seale L, Williams T, Ridgway A. An evaluation of stimulant medication on the reinforcing effects of play. *Journal of Applied Behavior Analysis* 2008;41(1):143-147.

57. Lundh A, Sismondo S, Lexchin J, Busuioc OA, Bero L. Industry sponsorship and research outcome. *Cochrane Database of Systematic Reviews* 2012(12):Art. No.: MR000033. DOI: 10.1002/14651858.MR000033.pub2.

58. Manos MJ, Short EJ, Findling RL. Differential effectiveness of methylphenidate and Adderall in school-age youths with attention-deficit/hyperactivity disorder. *Journal of the American Academy of Child and Adolescent Psychiatry* 1999;38(7):813-9.

59. March L, Irwig L, Schwarz J, et al. N-of-1 trials comparing a nonsteroidal anti-inflammatory drug with paracetamol in osteoarthritis. *British Medical Journal* 1994;309:1041–1045; discussion 1045–1046.

60. Mayes SD, Calhoun SL, Bixler EO, Vgontzas AN, Mahr F, Hillwig-Garcia J, et al. ADHD subtypes and comorbid anxiety, depression and oppositional-defiant disorder: differences in sleep problems. *Journal of Pediatric Psychology* 2009;34(33):328-37.

61. McCracken JT, Biederman J, Greenhill LL, Swanson JM, McGough JJ, Spencer TJ, et al. Analog classroom assessment of a once-daily mixed amphetamine formulation, SLI381 (Adderall XR), in children with ADHD. *Journal of the American Academy of Child and Adolescent Psychiatry* 2003;42(6):673-83.

62. Miller A, Lee S, Raina P, Klassen A, Zupancic J, Olsen L. A review of therapies for attention-deficit/hyperactivity disorder. www.cadth.ca/en/products/health-technology-assessment/publication/20 (accessed 11 February 2014).

63. Moher D, Liberati A, Tetzlaff J, Altman DG, The PRISMA Group. *P*referred *R*eporting *I*tems for *S*ystematic Reviews and *M*eta-*A*nalyses: The PRISMA Statement. *PLoS Medicine* 2009;6(6):e1000097. doi:10.1371/journal.pmed1000097

64. Moher D, Pham B, Lawson ML, Klassen TP. The inclusion of reports of randomized trials published in languages other than English in systematic reviews. *Health Technology Assessment* 2003;7:1-90.

65. National Institute of Health. National Institutes of Health consensus development conference statement: diagnosis and treatment of attention-deficit/hyperactivity disorder (ADHD). *Journal of the American Academy of Child and Adolescent Psychiatry* 2000;39(2):182-97.

66. Neef NA, Bicard DF, Endo S, Coury DL, Aman MG. Evaluation of pharmacological treatment of impulsivity in children with attention deficit hyperactivity disorder. *Journal of Applied Behavior Analysis* 2005;38(2):135-146.

67. Nemzer ED, Arnold E, Votolato NA, McConnell H. Amino acid supplementation as therapy for attention deficit disorder. *Journal of the American Academy of Child Psychiatry* 1986;25(4):509-13.

68. Nikles CJ, Mitchell GK, Del Mar CB, et al. An N-of-1 trial service in clinical practice: testing the effectiveness of stimulants for attention deficit/hyperactivity disorder. *Pediatrics* 2006;117:2040–2046.

69. Paterson C. Measuring outcome in primary care: a patient-generated measure, MYMOP, compared to the SF-36 health survey. *British Medical Journal* 1996;312:1016-20.

70. Payton JB, Burkhart JE, Hersen M, Helsel WJ. Treatment of ADDH in mentally retarded children: A preliminary study. *American Academy of Child and Adolescent Psychiatry* 1989;28(5):761-767.

71. Perdices M, Tate RL. Single-subject designs as a tool for evidence-based clinical practice: Are they unrecognised and undervalued? *Neuropsychological Rehabilitation* 2009;19(6):904-27.

72. Plint AC, Moher D, Morrison A, Schulz K, Altman DG, Hill C, Gaboury I. Does the CONSORT checklist improve the quality of reports of randomized controlled trials? A systematic review. *Medical Journal of Australia* 2006;185(5):263-267.

73. Pliszka SR, Browne RG, Olvera RL, Wynne SK. A double-blind, placebo-controlled study of Adderall and methylphenidate in the treatment of attention-deficit/hyperactivity disorder. Journal of the American Academy of Child and Adolescent Psychiatry 2000;39(5):619-26.

74. Polanczyk G, Silva de Lima M, Horta BL, Biederman J, Rohde LA. Worldwide prevalence of ADHD. *American Journal of Psychiatry* 2007;4(164):942-8.

75. Pope JE, Prashker M, Anderson J. The efficacy and cost effectiveness of N-of-1 studies with diclofenac compared to standard treatment with nonsteroidal anti-inflammatory drugs in osteoarthritis. *Journal of Rheumatology* 2004;31:140–149.

76. Porrino LJ, Rapoport JL, Behar D, Ismond DR, Bunney Jr. WE. A naturalistic assessment of motor activity of hyperactive boys (Stimulant drug effects). *Archives of General Psychiatry* 1983;40:688-693.

77. Punja S, Eslick I, Duan N, Vohra S and the DEcIDE Methods Center N-of-1 Guidance Panel (Gabler N, Kaplan H, Kravitz R, Larson E, Pace W, Schmid C, Sim I). Design and Implementation of N-of-1 Trials: A user's guide. An Ethical Framework for N-of-1-Trials: Clinical Care, Quality Improvement, or Human Subject Research? Agency for Healthcare Research and Quality; AHRQ Publication No.13(14)-EHC122-EF. Rockville, MD. February 2014

78. Punja S, Shamseer L, Hartling L, Urichuk L, Vandermeer B, Nikles CJ, Vohra S. Amphetamines for attention deficit/hyperactivity disorder in children and adolescents. *Cochrane Database of Systematic Reviews* [in press].

79. Punja S, Xu D, Schmid CH, Hartling L, Urichuk L, Nikles CJ, Vohra S. Amphetamines and methylphenidate for pediatric ADHD: A systematic review and meta-analysis of N-of-1 evidence. *Journal of Clinical Epidemiology* [in press].

80. Reitman D, Hupp SDA, O'Callaghan PM, Gulley V, Northup J. The influence of a token economy and methylphenidate on attentive and disruptive behavior during sports with ADHD-diagnosed children. *Behavior Modification* 2001;25(2):305-323.

81. Review Manager (RevMan) [Computer program]. Version 5.1. Copenhagen: The Nordic Cochrane Centre, The Cochrane Collaboration, 2011.

82. Riley RD, Lambert PC, Abo-Zaid G. Meta-analysis of individual participant data: Rationale, conduct, and reporting. *British Medical Journal* 2010;340:c221.

83. Riley RD, Lambert PC, Staessen JA, et al. Meta-analysis of continuous outcomes combining individual patient data and aggregate data. Statistics in Medicine 2008;27:1870-93.

84. Ross M, Olson JM. An expectancy-attribution model of the effects of placebos. *Psychology Review*. 1981;88(5):408-37.

85. Rothwell PM. External validity of randomized controlled trials: "to whom do the results of this trial apply?" *The Lancet* 2005; 365(9453):82-93.

86. Rutherford BR, Marcus SM, Wang P, Sneed JR, Pelton G, Devanand D, Duan, Roose SP. A randomized, prospective pilot study of patient expectancy and antidepressent outcome. *Psychological Medicine.* 2013;43(5):975-82.

87. Sackett D, Rosenberg W, Gray J, Haynes R, Richardson W. Evidence based medicine: What it is and what it isn't. *British Medical Journal* 1996; 312(7023):71-2.

88. Sampson M, Shamseer L, Bukutu C, Barrowman N, Moher D, Vohra S. Systematic reviews of N-of-1 methods, analysis and meta-analysis. Cochrane Canada 8th Annual Symposium, Ottawa: May 19-20 2010.

89. Schacter HM, Pham B, King J, Langford S, Moher D. How efficacious and safe is short-acting methylphenidate for the treatment of attention-deficit disorder in children and adolescents? A meta-analysis. *Canadian Medical Association Journal* 2001;165(11):1475-1488.

90. Schmid CH, Brown EN. Bayesian hierarchical models. *Methods in Enzymology* 2000;321:305-330.

91. Schulz KF, Altman DG, Moher D, for the CONSORT Group. CONSORT 2010 Statement: updated guidelines for reporting parallel group randomised trials. *BMJ* 2010;340:c332.

92. Senn S. Suspended judgment: N-of-1 Trials. *Controlled Clinical Trials* 1993; 14(1):1–5.

93. Shamseer L, Sampson M, Bukutu C, Barrowman N, Altman D, Moher D, Vohra S. P.05.50. CONSORT Extension for N-of-1 trials (CENT) Guidelines. *BMC Complementary and Alternative Medicine* 2012; 12(Suppl 1): P410.

94. Sharp WS, Walter JM, Marsh WL, Ritchie GF, Hamburger SD, Castellanos FX. ADHD in girls: clinical comparability of a research sample. *Journal of the American Academy of Child and Adolescent Psychiatry* 1999;38(1):40-7.

95. Shekim WO, Bylund DB, Alexson J, Glaser RD, Jones SB, Hodges K, et al. Platelet MAO and measures of attention and impulsivity in boys with attention deficit disorder and hyperactivity. *Psychiatry Research* 1986;18(2):179-88.

96. Short EJ, Manos MJ, Findling RL, Schubel EA. A prospective study of stimulant response in preschool children: insights from ROC analyses. *Journal of the American Academy of Child and Adolescent Psychiatry* 2004;43(3):251-9.

97. Speltz ML, Varley CK, Peterson K, Beilke RL. Effects of dextroamphetamine and contingency management on a preschooler with ADHD and oppositional defiant disorder. *American Academy of Child and Adolescent Psychiatry* 1988;27(2):175-178.

98. Spencer TJ, Wilens TE, Biederman J, Weisler RH, Read SC, Pratt R. Efficacy and safety of mixed amphetamine salts extended release (Adderall XR) in the management of attention-deficit/hyperactivity disorder in adolescent patients: a 4-week, randomized, double-blind, placebo-controlled, parallel-group study. *Clinical Therapeutics* 2006;28(2):266-79.

99. Stewart LA, Parmar MK. Meta-analysis of the literature or of individual patient data: is there a difference? *The Lancet* 1993; 341:418–422.

100. Stewart LA, Tierney JF. To IPD or not to IPD? Advantages and disadvantages of systematic reviews using individual patient data. *Evaluation and Health Professions* 2002;25(1):76-97.

101. Swanson JM, Kinsbourne M, Nigg J, et al. Etiologic subtypes of attention-

102. deficit/hyperactivity disorder: brain imaging, molecular genetic and environmental factors and the dopamine hypothesis. *Neuropsychology Review* 2007;17(1):39-59

103. Swanson JM, Wigal S, Greenhill LL, Browne R, Waslik B, Lerner M, et al. Analog classroom assessment of Adderall (R) in children with ADHD. *Journal of the American Academy of Child and Adolescent Psychiatry* 1998;37(5):519-26.

104. The Medical Letter Online. Lisdexamfetamine dimesylate (Vyvanase) for ADHD. *The Medical Letter on Drugs and Therapeutics* 2007;49(1265):58-9.

105. Volkow ND, Wang GJ, Newcorn J, et al. Depressed dopamine activity in caudate and preliminary evidence of limbic involvement in adults with attention-deficit/hyperactivity disorder. *JAMA Psychiatry* 2007;64(8):932-940.

106. Waxmonsky J. A randomized, double-blind, placebo-controlled, crossover, laboratory classroom study to evaluate the safety and efficacy of d-amphetamine transdermal drug delivery system (d-ATS) compared to placebo in children and adolescents with ADHD. http://clinicaltrials.gov/ct2/show/record/NCT01711021 (accessed 11 February 2014).

107. Wigal SB, Kollins SH, Childress AC, Squires L, Brams M, Childress A, et al. A 13-hour laboratory school study of lisdexamfetamine dimesylate in school-aged children with attention-deficit/hyperactivity disorder. Child *and Adolescent Psychiatry and Mental Health* 2009;3(1):17.

108. Wigal T, Swanson JM, Regino R, Lerner MA, Soliman I, Steinoff K, et al. Stimulant medications for the treatment of ADHD: efficacy and limitations. Mental Retardation and Development Disabilities Research Reviews 1999;5(3):215-24.

109. Zarin DA, Young JL, West JC. Challenges to evidence-based medicine. *Social Psychiatry and Psychiatric Epidemiology* 2005;40(1):27-35.

110. Zucker DR, Ruthazer R, Schmid CH, et al. Lessons learned combining N-of-1 trials to assess fibromyalgia therapies. *Journal of Rheumatology* 2006;33:2069–2077.

111. Zucker DR, Ruthazer R, Schmid CH. Individual (N-of-1) trials can be combined to give population comparative treatment effect estimates: methodologic considerations. *Journal of Clinical Epidemiology* Dec 2010;63(12):1312-1323

112. Zucker DR, Schmid CH, McIntosh MW, D'Agostino RB, Selker HP, Lau J. Combining single patient (N-of-1) trials to estimate population treatment effects and to evaluate individual patient responses to treatment. *Journal of Clinical Epidemiology* 1997 04;50(4):401-410.

113. Zwaigenbaum L, Dick P, Handley-Derry M, Malone M, Jacobson S. "N-of-1" trials of methylphenidate in two children with Williams Syndrome and attention deficit hyperactivity disorder. *Journal of Developmental and Physical Disabilities* 2006;18(1):45-56.