The purpose of this report is to make available information regarding the comparative effectiveness and safety profiles of different drugs within pharmaceutical classes. Reports are not usage guidelines, nor should they be read as an endorsement of, or recommendation for, any particular drug, use or approach. Oregon Health & Science University does not recommend or endorse any guideline or recommendation developed by users of these reports.
**TABLE OF CONTENTS**

**INTRODUCTION** .............................................................................................................................. 4  
Scope and Key Questions ....................................................................................................................... 6

**METHODS** ...................................................................................................................................... 8  
Literature Search ..................................................................................................................................... 8  
Study Selection ....................................................................................................................................... 9  
Data Abstraction .................................................................................................................................... 9  
Quality Assessment ................................................................................................................................. 9  
Evidence Synthesis ................................................................................................................................ 10

**RESULTS** ........................................................................................................................................ 11  
Overview ............................................................................................................................................... 11  
Overall Summary .................................................................................................................................. 13  
Key Question 1: Efficacy and effectiveness .......................................................................................... 20  
Key Question 2: Tolerability and safety ............................................................................................... 38  
Key Question 3: Subgroups .................................................................................................................. 48

**OVERALL SUMMARY** .................................................................................................................. 56

**REFERENCES** .................................................................................................................................. 61

**In-text Tables**
- Table 1. ADHD drugs and doses .................................................................................................... 7  
- Table 2. Numbers of Head to Head Trials of Drugs for ADHD ......................................................... 12  
- Table 3. Dextroamphetamine IR versus Methylphenidate IR Study Characteristics ....................... 24  
- Table 4. Comparison of Response Rates to MPH IR .................................................................... 25  
- Table 5. Maintenance of MPH IR short-term effects .................................................................... 27  
- Table 6. Summary of Differences in Results of ADHD/Tourette’s Disorder Study ....................... 31  
- Table 7. Adult ADHD – Comparative trial outcomes ..................................................................... 34  
- Table 8. Adult ADHD – Response Rates from Placebo-Controlled Trials .................................... 35  
- Table 9. Adult ADHD – Other symptom-related outcomes in PCTs ............................................. 36  
- Table 10. Long-term functional outcomes of MPH from Hechtman 1984 .................................... 37  
- Table 11. Summary of Adverse Effects Reported ....................................................................... 39  
- Table 12. Specific Adverse Events in placebo-controlled trials of adults ...................................... 41  
- Table 13. Long-Term Height and Weight Outcomes in Observational Studies ............................ 43  
- Table 14. Tic-related outcomes in observational studies ............................................................... 46  
- Table 15. Tic Severity Ratings in trials of drugs to treat ADHD ..................................................... 51  
- Table 16. Main Outcomes in Studies in Children with ADHD and Mental Retardation .......... 52  
- Table 17. Adverse Events in Studies in Children with ADHD and Mental Retardation ............. 53  
- Table 18. OVERALL SUMMARY TABLE ................................................................................... 56

**Figures**
- Figure 1. ADHD Drug Class Review Flow Diagram ..................................................................... 74
Appendices

Appendix A. Scales used to assess efficacy and adverse events .............................................75
Appendix B. Search strategy ...................................................................................................82
Appendix C. Methods for drug class reviews ........................................................................88
Appendix D. Excluded studies .................................................................................................92
Appendix E. Previous systematic reviews .............................................................................113

Evidence Tables – published in a separate document.
INTRODUCTION

According to the most recent NIH Consensus Statement (1998), attention deficit hyperactivity disorder (ADHD) is the most commonly diagnosed childhood behavioral disorder.”\(^1\) Classification of hyperactivity and defects in attention emerged in the 1960’s as Minimal Brain Dysfunction (MBD) and Hyperkinetic Syndrome, and has continued to evolve over time.\(^2\)

A number of community-based studies have reported ADHD prevalence rates that range from 1.7% to 16%.\(^3\) This is broader than the range of 3 to 5 percent that was estimated by the expert panelists that participated in the NIH Consensus Development Conference on Diagnosis and Treatment of Attention Deficit Hyperactivity Disorder (ADHD) in 1998.\(^1\) The estimated prevalence cited in the most recent (1997) version of the Diagnostic and Statistical Manual of Mental Disorders (DSM-IV) is 3 to 7 percent.\(^4\) Differences in prevalence estimates may be due to variation in methods of ascertainment and diagnostic criteria.\(^5\) While no independent diagnostic test exists for ADHD, the DSM-IV provides standardized criteria that can be used as a foundation for clinical diagnosis.\(^1, 4\) According to the DSM-IV, essential features of ADHD include persistent levels of inattention, impulsivity and/or hyperactivity that exceed usual developmental patterns.\(^4\) In order to qualify for a DSM-IV diagnosis of ADHD, symptoms must date back to before age 7, persist for at least six months, and cause impairment that interferes with functional capacity in at least two performance settings (social, academic, or employment).\(^4\) DSM-IV specifies three distinct subtypes of ADHD that are characterized by predominantly inattentive, hyperactive-impulsive, or mixed symptoms.\(^4\)

ADHD is diagnosed more frequently in males than in females.\(^6\) Comorbidities such as mood and conduct disorders, tics or Tourette syndrome, learning disorders and mental retardation may be found in up to 65% of individuals with ADHD.\(^3\) With regard to the course of ADHD, symptoms can persist into adolescence in 80 percent of cases and into adulthood in 65 percent of cases.\(^6\)

Drugs Used to Treat ADHD:
Drug therapy of ADHD has consisted primarily of stimulant medications, but more recently nonstimulant medications have been investigated because of concerns over the long term effects of stimulants on growing children. Throughout the report dextroamphetamine with be abbreviated to DEX and methylphenidate will be abbreviated to MPH.

Stimulants
Amphetamine mixture (Amphetamine mixture): Amphetamines are non-catecholamine sympathomimetic amines with CNS stimulant activity. Dextroamphetamine sulfate is the dextro isomer of the compound \(d,l\) -amphetamine sulfate, a sympathomimetic amine of the amphetamine group.

Methylphenidate HCL is a mild central nervous system stimulant. The mode of action in man is not completely understood, but it presumably activates the brain stem arousal system and cortex to produce its stimulant effect. Dextymethylphenidate HCL is the more pharmacologically active enantiomer of the \(d\) - and \(l\) –enantiomers of methylphenidate, is thought to block the reuptake of norepinephrine and dopamine into the presynaptic neuron and increase the release of these monoamines into the extraneuronal space.
Pemoline is a central nervous system stimulant. Pemoline has a pharmacological activity similar to that of other known central nervous system stimulants; however, it has minimal sympathomimetic effects and is structurally dissimilar to the amphetamines and methylphenidate. Although studies indicate that pemoline may act in animals through dopaminergic mechanisms, the exact mechanism and site of action of the drug in man is not known. There is neither specific evidence which clearly establishes the mechanism whereby pemoline produces its mental and behavioral effects in children, nor conclusive evidence regarding how these effects relate to the condition of the central nervous system. Because of potential for liver toxicity, regular monitoring of liver function is required and pemoline is reserved for second-line therapy of ADHD.

Modafinil: Modafinil is a central nervous system stimulant approved for promoting wakefulness, although the precise mechanism(s) is unknown. Modafinil has wake-promoting actions like sympathomimetic agents including amphetamine and methylphenidate, although the pharmacologic profile is not identical to that of sympathomimetic amines. At pharmacologically relevant concentrations, modafinil does not bind to most potentially relevant receptors for sleep/wake regulation, including those for norepinephrine, serotonin, dopamine, GABA, adenosine, histamine-3, melatonin, or benzodiazepines. Modafinil also does not inhibit the activities of MAO-B or phosphodiesterases II-V.

Non-stimulants
Atomoxetine HCl: The precise mechanism by which atomoxetine produces its therapeutic effects in ADHD is unknown, but is thought to be related to selective inhibition of the pre-synaptic norepinephrine transporter, as determined in ex vivo uptake and neurotransmitter depletion studies.

Antidepressants. Bupropion HCl: The neurochemical mechanism of the antidepressant effect of bupropion is not known. Bupropion is a relatively weak inhibitor of the neuronal uptake of norepinephrine, serotonin, and dopamine, and does not inhibit monoamine oxidase. Bupropion produces dose-related central nervous system stimulant effects in animals, as evidenced by increased locomotor activity, increased rates of responding in various schedule-controlled operant behavior tasks, and, at high doses, induction of mild stereotyped behavior. Bupropion is not currently approved for use in ADHD.

Antihypertensives. Clonidine HCl stimulates alpha-adrenoreceptors in the brain stem. This action results in reduced sympathetic outflow from the central nervous system and in decreases in peripheral resistance, renal vascular resistance, heart rate, and blood pressure. Other studies in patients have provided evidence of a reduction in plasma renin activity and in the excretion of aldosterone and catecholamines. Guanfacine HCl is a centrally acting alpha-adrenergic agonist with similar actions to clonidine, but is thought to cause less sedation. Clonidine and guanfacine are not currently approved for use in ADHD. Clonidine and guanfacine acutely stimulate growth hormone release in both children and adults, but do not produce a chronic elevation of growth hormone with long-term use. The mechanism of action for effects seen with clonidine or guanfacine in treating ADHD is not known.
Atypical antipsychotic drugs (Aripiprazole, Clozapine, Olanzapine, Quetiapine, Risperidone, Ziprasidone) have varying mechanisms of action, but all involve antagonism of subsets of dopamine and serotonin receptors. None of the atypical antipsychotic drugs are currently approved for use in ADHD.

Scope and Key Questions

The purpose of this review is to compare the benefits and harms of different pharmacologic treatments for ADHD. The Oregon Evidence-based Practice Center wrote preliminary key questions, identifying the populations, interventions, and outcomes of interest, and based on these, the eligibility criteria for studies. These were reviewed and revised by representatives of organizations participating in the Drug Effectiveness Review Project (DERP). The participating organizations of DERP are responsible for ensuring that the scope of the review reflects the populations, drugs, and outcome measures of interest to both clinicians and patients. The participating organizations approved the following key questions to guide this review:

**Key Question 1.** What is the comparative effectiveness of different pharmacologic treatments for attention deficit disorders?

**Key Question 2.** What is the comparative tolerability and safety of different pharmacologic treatments for attention deficit disorders?

**Key Question 3.** Are there subgroups of patients based on demographics (age, racial groups, gender), other medications, or co-morbidities for which one pharmacologic treatment is more effective or associated with fewer adverse events?

Inclusion Criteria

**Populations**

Pediatric and adult outpatients with Attention Deficit Disorders

- Attention Deficit Disorder
- Attention Deficit Hyperactivity Disorder
Interventions (immediate release and extended release formulations, where applicable)

Table 1. ADHD drugs and doses

<table>
<thead>
<tr>
<th>Generic Name</th>
<th>Trade Name*</th>
<th>FDA ADHD Approval</th>
<th>Year Introduced</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Stimulants</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Amphetamine mixture**</td>
<td>Adderall®*</td>
<td>Children</td>
<td>1960</td>
</tr>
<tr>
<td></td>
<td>Adderall XR***</td>
<td>Both</td>
<td>2001</td>
</tr>
<tr>
<td>Dextroamphetamine sulfate</td>
<td>Dextedrine®*</td>
<td>Children</td>
<td>1976</td>
</tr>
<tr>
<td></td>
<td>Dextrostat®*</td>
<td>Children</td>
<td>1975</td>
</tr>
<tr>
<td></td>
<td>Dextedrine Spansule®</td>
<td>Children</td>
<td>Unknown</td>
</tr>
<tr>
<td>Methylphenidate HCL</td>
<td>Focalin®</td>
<td>Children</td>
<td>2001</td>
</tr>
<tr>
<td>Metadate CD® (MPH OROS)</td>
<td>Children</td>
<td>2001</td>
<td></td>
</tr>
<tr>
<td>Metadate ER® (MPH ER)</td>
<td>Both</td>
<td>1999</td>
<td></td>
</tr>
<tr>
<td>Methylin®</td>
<td>Both</td>
<td>2003</td>
<td></td>
</tr>
<tr>
<td>Ritalin®*</td>
<td>Both</td>
<td>1955</td>
<td></td>
</tr>
<tr>
<td>Ritalin SR® (MPH SR)</td>
<td>Both</td>
<td>1982</td>
<td></td>
</tr>
<tr>
<td>Ritalin LA® (MPH SODAS)</td>
<td>Children</td>
<td>2002</td>
<td></td>
</tr>
<tr>
<td><strong>Modafinil</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Pemoline</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Non-Stimulants</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Atomoxetine HCl</td>
<td>Strattera®</td>
<td>Both</td>
<td>2002</td>
</tr>
<tr>
<td><strong>Antidepressants</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Buproprion HCl</td>
<td>Wellbutrin®</td>
<td>None</td>
<td>1985</td>
</tr>
<tr>
<td></td>
<td>Wellbutrin SR®</td>
<td>None</td>
<td>1996</td>
</tr>
<tr>
<td></td>
<td>Wellbutrin XL®</td>
<td>None</td>
<td>2003</td>
</tr>
<tr>
<td><strong>Antihypertensives</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Clonidine HCl</td>
<td>Catapres®*</td>
<td>None</td>
<td>1974</td>
</tr>
<tr>
<td>Guanfacine HCl</td>
<td>Tenex®*</td>
<td>None</td>
<td>1986</td>
</tr>
<tr>
<td><strong>Atypical Antipsychotics</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Aripiprazole</td>
<td>Abilify®</td>
<td>None</td>
<td>2002</td>
</tr>
<tr>
<td>Clozapine</td>
<td>Clozaril®*</td>
<td>None</td>
<td>1989</td>
</tr>
<tr>
<td>Olanzapine</td>
<td>Zyprexa®</td>
<td>None</td>
<td>1996</td>
</tr>
<tr>
<td>Quetiapine</td>
<td>Seroquel®</td>
<td>None</td>
<td>1997</td>
</tr>
<tr>
<td>Risperidone</td>
<td>Risperdal®</td>
<td>None</td>
<td>1993</td>
</tr>
<tr>
<td>Ziprasidone</td>
<td>Geodon®</td>
<td>None</td>
<td>2001</td>
</tr>
</tbody>
</table>

*or generic equivalent
** (amphetamine aspartate; amphetamine sulfate; dextroamphetamine saccharate; dextroamphetamine sulfate)
***no longer marketed in Canada (February 2005)

Outcomes
- Symptom response (inattention, hyperactivity-impulsivity, aggression, global ratings, etc.)
- Functional capacity (social, academic and occupational productivity)
- Caregiver satisfaction (parent, teacher)
- Quality of life (child, parent, caregivers, teachers)
- Overall adverse effect reports
- Withdrawals due to adverse effects
- Serious adverse events reported
- Specific adverse events (hepatotoxicity, insomnia, anorexia, effects on growth, abuse potential)
• Time to onset of effectiveness
• Duration of effectiveness

Scales and Tests Used to Measure Outcomes
Numerous ADHD-specific and other psychiatric rating scales, as well as neuropsychological testing methods are used to measure symptoms of ADHD. We limited our analyses to rating scales/tests for which we found published evidence of good reliability and validity. Our primary sources for documentation of the psychometric properties of rating scales included the Agency for Healthcare Research and Quality (AHRQ) Technical Review #3 (Diagnosis of Attention-Deficit/Hyperactivity Disorder)7 and Mental Measurements Yearbooks.8-15 The AHRQ Technical Review #3 provides qualitative information on many of the rating scales cited in our report, including “subscales included in each test, comorbid conditions addressed by each checklist, time required to administer, number of items, ages for which norms are available, computer scoring availability, and ordering information, including cost” and reliability and validity. Appendix A provides a listing of commonly used scales and tests and associated acronyms.

Study designs
• Controlled clinical trials and good-quality systematic reviews.
• Observational studies with functional or adverse event outcomes

The benefit of the RCT design is the ability to a reliably unbiased estimate of treatment effects in a controlled setting, by randomizing patients is the best method of producing comparable groups based on both known and unknown prognostic factors.16,17 However, RCTs can vary in quality, and often suffer from limitations in generalizability to the larger patient population. Observational study designs are thought to have greater risk of introducing bias, although they typically represent effects in a broader section of the overall patient population. While it has been shown that some observational studies and RCTs of the same treatments have similar findings, there are also multiple example of situations where this has not been true and the question of what type of evidence is best has not been resolved.18,19 While RCTs also provide good evidence on short-term adverse events, observational designs are useful in identifying rare, serious adverse events which often require large numbers of patients exposed to a treatment over longer periods of time to be identified.

METHODS

Literature Search
To identify relevant citations, we searched the Cochrane Central Register of Controlled Trials (2nd Quarter 2004), Cochrane Database of Systematic Reviews, MEDLINE (1966 to June Week 2 2004), EMBASE (2nd Quarter 2004), and PsycINFO (1974 to May Week 5 2004) using terms for included drugs, indications, and study designs (see Appendix B for complete search strategies). We have attempted to identify additional studies through searches of reference lists of included studies and reviews, the FDA web site, as well as searching dossiers submitted by pharmaceutical companies for the current review. All citations were imported into an electronic database (EndNote 6.0).
Study Selection
We assessed abstracts of citations identified from literature searches for inclusion, using the criteria described above. Full-text articles of potentially relevant abstracts were retrieved and a second review for inclusion was conducted by reapplying the inclusion criteria.

Data Abstraction
The following data were abstracted from included trials: study design, setting, population characteristics (including sex, age, ethnicity, diagnosis), eligibility and exclusion criteria, interventions (dose and duration), comparisons, numbers screened, eligible, enrolled, and lost to follow-up, method of outcome ascertainment, and results for each outcome. We recorded intention-to-treat results when reported. If true intention-to-treat results were not reported, but loss to follow-up was very small, we considered these results to be intention-to-treat results. In cases where only per-protocol results were reported, we calculated intention-to-treat results if the data for these calculations were available. In trials with crossover, outcomes for the first intervention were recorded if available. This was because of the potential for differential withdrawal prior to crossover biasing subsequent results and the possibility of either a “carryover effect” (from the first treatment) in studies without a washout period, or “rebound” effect from withdrawal of the first intervention.

Data abstracted from observational studies included design, eligibility criteria duration, interventions, concomitant medication, assessment techniques, age, gender, ethnicity, number of patients screened, eligible, enrolled, withdrawn, or lost to follow-up, number analyzed, and results.

Quality Assessment
We assessed the internal validity (quality) of trials based on the predefined criteria listed in Appendix C. These criteria are based on the U.S. Preventive Services Task Force and the National Health Service Centre for Reviews and Dissemination (U.K.) criteria.\textsuperscript{20, 21} We rated the internal validity of each trial based on the methods used for randomization, allocation concealment, and blinding; the similarity of compared groups at baseline; maintenance of comparable groups; adequate reporting of dropouts, attrition, crossover, adherence, and contamination; loss to follow-up; and the use of intention-to-treat analysis. Trials that had a fatal flaw in one or more categories were rated “poor-quality”; trials that met all criteria were rated “good-quality”; the remainder were rated “fair-quality.” A fatal flaw occurs when there is evidence of bias or confounding in the trial, for example when randomization and concealment of allocation of random order are not reported and baseline characteristics differ significantly between the groups. In this case, randomization has apparently failed and for one reason or another bias has been introduced.

As the fair-quality category is broad, studies with this rating vary in their strengths and weaknesses: the results of some fair-quality studies are likely to be valid, while others are only probably valid. A poor-quality trial is not valid—the results are at least as likely to reflect flaws in the study design as the true difference between the compared drugs. External validity of trials was assessed based on whether the publication adequately described the study population, how similar patients were to the target population in whom the intervention will be applied, and whether the treatment received by the control group was reasonably representative of standard practice. We also recorded the role of the funding source.
Appendix C also shows the criteria we used to rate observational studies. These criteria reflect aspects of the study design that are particularly important for assessing adverse event rates. We rated observational studies as good-quality for adverse event assessment if they adequately met six or more of the seven predefined criteria, fair-quality if they met three to five criteria and poor-quality if they met two or fewer criteria.

Included systematic reviews were also rated for quality based on pre-defined criteria (see Appendix C), based on a clear statement of the questions(s), inclusion criteria, adequacy of search strategy, validity assessment and adequacy of detail provided for included studies, and appropriateness of the methods of synthesis.

Overall quality ratings for the individual study were based on internal and external validity ratings for that trial. A particular randomized trial might receive two different ratings: one for effectiveness and another for adverse events. The overall strength of evidence for a particular key question reflects the quality, consistency, and power of the set of studies relevant to the question.

Evidence Synthesis

Effectiveness versus Efficacy

Throughout this report, we highlight effectiveness studies conducted in primary care or office-based settings that use less stringent eligibility criteria, assess health outcomes, and have longer follow-up periods than most efficacy studies. Effectiveness studies conducted in primary care or office-based settings use less stringent eligibility criteria, assess health outcomes, and have longer follow-up periods than most efficacy studies. The results of effectiveness studies are more applicable to the “average” patient than results from highly selected populations in efficacy studies. Examples of “effectiveness” outcomes include quality of life, global measures of academic success, and the ability to work or function in social activities. These outcomes are more important to patients, family and care providers than surrogate or intermediate measures such as scores based on psychometric scales.

An evidence report pays particular attention to the generalizability of efficacy studies performed in controlled or academic settings. Efficacy studies provide the best information about how a drug performs in a controlled setting that allow for better control over potential confounding factors and bias. However, the results of efficacy studies are not always applicable to many, or to most, patients seen in everyday practice. This is because most efficacy studies use strict eligibility criteria which may exclude patients based on their age, sex, medication compliance, or severity of illness. For many drug classes severely impaired patients are often excluded from trials. Often, efficacy studies also exclude patients who have “comorbid” diseases, meaning diseases other than the one under study. Efficacy studies may also use dosing regimens and follow up protocols that may be impractical in other practice settings. They often restrict options, such as combining therapies or switching drugs, that are of value in actual practice. They often examine the short-term effects of drugs that, in practice, are used for much longer periods of time. Finally, they tend to use objective measures of effect that do not capture all of the benefits and harms of a drug or do not reflect the outcomes that are most important to patients and their families.
Data Presentation

We constructed evidence tables showing the study characteristics, quality ratings, and results for all included studies. Studies that evaluated one pharmacologic treatment of ADHD against another provided direct evidence of comparative benefits and harms. Outcomes of changes in symptom measured using scales or tools with good validity and reliability are preferred over scales or tools with low validity/reliability or no reports of validity/reliability testing. Where possible, head-to-head data are the primary focus of the synthesis.

In theory, trials that compare these drugs to other interventions or placebos can also provide evidence about effectiveness. This is known as an indirect comparison and can be difficult to interpret for a number of reasons, primarily issues of heterogeneity between trial populations, interventions, and assessment of outcomes. Indirect data are used to support direct comparisons, where they exist, and are also used as the primary comparison where no direct comparisons exist. Such indirect comparisons should be interpreted with caution.

Throughout the report dextroamphetamine with be abbreviated to DEX and methylphenidate will be abbreviated to MPH.

RESULTS

Overview

We identified 2,287 citations from literature searches and reviews of reference lists. This includes citations from dossiers submitted by six pharmaceutical manufacturers: Shire US (amphetamine mixture, Amphetamine mixture XL), Eli Lilly (atomoxetine HCl), GlaxoSmithKline (bupropion HCl), McNeil (methylphenidate HCL, Concerta®), Novartis (methylphenidate HCL, Ritalin LA®), and Cephalon (modafinil). After applying the eligibility and exclusion criteria to the titles and abstracts, we obtained full-paper copies of 514 publications. After re-applying the criteria for inclusion, we ultimately included 180 publications (146 studies and 43 duplicate data or background publications). A list of excluded studies is reported in Appendix D. The flow of study inclusion and exclusion is detailed in Figure 1.

We identified the following numbers of head to head comparative trials of pharmacologic treatments for ADHD.
### Table 2. Numbers of Head to Head Trials of Drugs for ADHD

<table>
<thead>
<tr>
<th></th>
<th>MPH IR</th>
<th>MPH ER</th>
<th>DEX</th>
<th>DEX-MPH</th>
<th>Adderall</th>
<th>Modafinil</th>
<th>Pemoline</th>
<th>Atomoxetine</th>
<th>Bupropion</th>
<th>Clonidine</th>
<th>Guanfacine</th>
<th>AAPs</th>
</tr>
</thead>
<tbody>
<tr>
<td>MPH IR</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>MPH ER</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>DEX</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>DEX-MPH</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Adderall</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Modafinil</td>
<td>--</td>
<td>--</td>
<td>A: 1</td>
<td>--</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pemoline</td>
<td></td>
<td></td>
<td></td>
<td>C: 1</td>
<td>--</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Atomoxetine</td>
<td>C: 5</td>
<td>C: 1</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Bupropion</td>
<td>C: 1</td>
<td>A: 1</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Clonidine</td>
<td>C: 3</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Guanfacine</td>
<td>--</td>
<td>--</td>
<td>A: 1</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>AAPs</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

C= children  
T= adolescents  
A= adults

Data abstracted from these trials can be found in Evidence Table 3 and 9 and the relevant quality assessments in Evidence Table 4 and 10. We found hundreds of placebo-controlled trials in children. The majority were studies of MPH IR and fewer of various other formulations (OROS, SR and MR). Placebo-controlled trials also studied DEX, clonidine, pemoline, bupropion, Adderall, unspecified “psychostimulants”, atomoxetine, and guanfacine, modafinil, “amphetamine” and risperidone. Because there are a large number of head to head trials, and indirect comparisons from placebo controlled trials are less reliable, we have included 4 placebo-controlled trials of drugs for which we have no head to head evidence (dexamethylenphidate, guanfacine, MPH ER, MPH CD or modafinil) and eight for which we have only one head to head trial (atomoxetine, bupropion, and clonidine). Also for Key Question 1, we included 6 placebo-controlled trials and 3 multimodal trials and 5 observational studies as the only evidence of remission rates and long-term functional outcomes and response maintenance. One trial of the effects on weight and height of children discontinuing MPH during summers was included (MPH versus no treatment) was included in Key Question 2. We also included 16 placebo-controlled trials in subgroup populations in Key Question 3. Data abstracted from placebo controlled trials can be found in Evidence Table 5, and relevant quality assessments in Evidence Table 6. For long-term safety, we included 18 observational studies (Evidence Tables 15 and 16).

In adult populations (age 18 and above), we included 8 placebo-controlled trials (Evidence Tables 11 and 12), and one long-term observational study (Evidence Tables 15 and 16) in addition to the head to head trials listed in Table 2 above.
Previous Systematic Review Findings

Three good quality systematic reviews of drug treatments for ADHD were conducted prior to this review, one in the US\textsuperscript{5}, one in Canada\textsuperscript{5}, and one in the UK\textsuperscript{22}, were included. There are some differences in the lists of drugs reviewed amongst these reviews and our report, with the commonalities being MPH IR and SR formulations, DEX, atomoxetine, pemoline, bupropion, and clonidine. The Canadian and British reviews did not included adults. These reviews consistently found a lack of evidence of a difference between the drugs studied in efficacy or adverse events. In some part, the reason for not finding a difference was thought to be due to small sample sizes lacking power to find a difference, and some studies were given less weight due to poor quality. Differences in adverse events were thought to be minor, although the assessment and reporting of adverse events was criticized. These reviewers also commented on the lack of good quality studies assessing long-term outcomes, both of effectiveness and serious adverse events. See Appendix E for further description of the findings of these reviews.

The American Academy of Pediatrics Clinical Practice Guideline on treatment of school-aged children with ADHD was also reviewed.\textsuperscript{23} The guideline considers only stimulant medications, specifically all forms of MPH and DEX. Stimulant and/or behavior therapy is recommended, the guideline does not prefer one, and states that the Jadad review (cited above) found no difference between these stimulants. The guideline also states, “Individual children, however, may respond to one of the stimulants but not to another.”

What this review adds

Our review adds to these prior reviews in a number of areas. Cross-referencing lists of studies included in each review reveals that we have included several studies that the other reviews did not.\textsuperscript{24-47} Some of the reasons for these studies not being included in the other reviews are differences in the scope of drugs reviewed, the outcomes included, and study designs included. For example, our review included stimulants (Adderall, modafinil), and non-stimulants (atypical antipsychotics, clonidine and guanfacine) not included in these reviews. Importantly, the current review includes observational studies to assess harms and functional outcomes and RCTs with functional outcomes such as academic achievement that were not included in these previous reviews. This review includes comparative evidence on the effect of MPH on weight and height, which was not included in the previous reviews. In addition, special effort has been made to identify the effects of ADHD subtype, diagnostic tool or definition, primary outcomes, comorbidities and ethnicity.

Overall Summary of the Evidence on Efficacy or Effectiveness, Short-Term Efficacy and Tolerability, and Long-Term Safety of Drugs Used to Treat ADHD.

General

- There are no trials of comparative effectiveness of these drugs for treatment of ADHD.
- Good quality evidence on the use of drugs to affect outcomes relating to global academic performance, consequences of risky behaviors, social achievements, etc. is lacking.
- The evidence for comparative efficacy and adverse events of drugs for treating ADHD is severely limited by small sample sizes, very short durations, and the lack of studies measuring functional or long-term outcomes. Methods of measuring symptom control vary significantly across studies. The crossover design was frequently used, with few analyzing the effect of order of administration of drugs and those that did found a
significant effect. No head-to-head efficacy trial was good quality. The small numbers of patients in these trials limits the ability to show a difference between drugs if one exists.

- Limitations to the generalizability of these trials include the following. The small sample sizes of these trials did not allow for statistical analyses of potential effects of these factors
  - Variation in diagnostic criteria and/or level of severity required for enrollment
  - Proportions of patients with various subtypes of ADHD
  - Proportions of patients with various comorbidities
  - Lack of inclusion of a variety of ethnic groups

- Overall, the rate of response to stimulants appears to be in the range of 60 to 80%, however the definitions of response rate varied and may not be comparable. Depending on the definition used, there is lack of clarity on the relationship of response rate to clinical significance. Response rates of nonstimulants vary, but the range in placebo-controlled trials is similar to that found with stimulants. Significant variation in the method of assessment and definition of response are most likely the reason for the wide variation.

**Young Children (Preschool Age; 3-5 years)**

**Efficacy and tolerability**

- No comparative evidence in young children was found.
- MPH was superior to placebo in efficacy in only 1 of 3 fair-quality placebo-controlled trials that used a validated assessment tool (CPRS-R); but was also associated with higher rates of adverse events

**Long-Term Safety**

- We found no evidence on long-term safety of drugs used to treat ADHD in young children

**Children (Elementary School Age; 6-12 years)**

**Effectiveness**

- Because no trials of effectiveness were found, observational studies were assessed for outcomes of effectiveness.
- The only comparative study with relevant outcomes found MPH OROS to be associated with fewer outpatient visits/hospitalization for accidents/injury than MPH IR over 12-months. Methodologic concerns over this study suggest caution in interpretation of these findings.
- Uncontrolled observational data assessing the effect of duration of treatment with MPH IR found no differences in academic achievement as measured by teachers, the proportion repeating grades, in special education classes or being tutored. Again, significant methodologic limitations suggest caution in interpreting these findings.

**Efficacy and tolerability**

**Stimulants**

- Immediate Release versus Extended Release formulations
  - Including the largest RCT, with 312 children, 2 studies of MPH IR versus MPH OROS (Concerta®) did not show an overall difference in outcomes. Very limited evidence was mixed in the comparison on MPH IR and MPH SR (Ritalin SR®).
Database studies using intermediate outcomes report greater persistence with MPH OROS and MPH SODAS compared to MPH IR. Methodologic concerns indicate caution in interpreting this evidence.

**Sustained Release versus Sustained Release**

- Very limited evidence suggests that MPH XR (Ritalin LA®) was superior to MPH OROS (Concerta®) on some, but not all efficacy outcomes. However, these results should be interpreted with caution until further evidence is available. We did not find evidence of a difference in adverse events between IR and SR formulations.

**Dextroamphetamine versus Methylphenidate**

- The body of evidence clearly indicates no difference in efficacy between DEX and MPH IR. Evidence from short-term trials and observational studies suggests that weight loss is greater with DEX than MPH IR.

**Adderall® versus Methylphenidate**

- Amphetamine mixture was superior to MPH IR on a few efficacy outcome measures in two trials, but clear evidence of superiority is lacking. Very limited evidence suggests that twice daily dosing of Amphetamine mixture led to higher rates of loss of appetite and sleep trouble than once daily dosing or MPH IR.

**Dextroamphetamine versus Adderall®**

- Evidence on the comparison of DEX IR versus SR versus Amphetamine mixture is limited and conflicting, but may suggest that measures made in the morning show DEX IR superior to DEX SR, and afternoon measures show DEX SR superior to Amphetamine mixture. Transient weight loss was greater with Amphetamine mixture and DEX SR than with DEX IR. However, this evidence should be interpreted with caution.

**Other stimulants**

- Head-to-head studies of Pemoline were poor quality and no conclusions can be drawn.

**Longer-term trials**

- Longer-term trials of MPH IR, placebo or non-medication treatments provide some evidence to assess the ability of MPH IR to maintain effects longer periods of time. These trials report somewhat mixed results on the ability to maintain short-term improvements in symptoms over 6 to 24 months. While the 14-month MTA found no deterioration over time, 3 other studies found the reverse. One explanation for this finding may be dose. One study found that the higher dose groups did not have deterioration of the gains in symptom control of inattentiveness and hyperactivity, and found an overall dose-response. The mean dose in the MTA study was also higher that in these other trials, and used 3 doses per day. A 10 month follow up of patients from the MTA study showed a decrease in the magnitude of effect.

**Nonstimulants**

**Atomoxetine**

- Limited evidence suggests a lack of a difference in efficacy. Atomoxetine caused more vomiting and somnolence than MPH IR, while MPH IR caused more ‘abnormal thinking’.

**Bupropion**

- Limited evidence suggests a lack of a difference in efficacy. No difference between adverse events reported with bupropion or MPH IR was found.
• Clonidine
  o There is some evidence that clonidine was superior to MPH IR in improving tic outcome measures, while few differences were found on ADHD measures. Limited head-to-head evidence suggests clonidine causes a higher incidence and greater degree of sedation than MPH IR.

• Other non-stimulants
  o While no head-to-head evidence is available, placebo-controlled trials of modafinil and guanfacine indicated that each drug was superior to placebo on most efficacy measures. The guanfacine study enrolled children comorbid with tic disorders, and also found improvement in tics.
  o No evidence was found for the atypical antipsychotics aripiprazole, clozapine, quetiapine, and ziprasidone. Indirect evidence of an effect of risperidone compared to placebo comes from a subgroup analysis of children with ADHD in a larger trial, indicating a beneficial effect. Only incomplete or poor quality evidence was found for dexamethasone, which is FDA approved for ADHD.

Long-Term Safety
• Although the observational studies provide some estimate of the prevalence of serious longer-term adverse events with amphetamine mixture, atomoxetine, DEX, and MPH (IR and SR), few studies directly compared different pharmacologic treatments for ADHD for any one adverse event.
• For outcomes where only uncontrolled evidence is available, it is not possible to draw conclusions about comparative long-term safety through indirect comparisons across observational studies due to large differences in study characteristics.
• The overall body of evidence is poor quality due to a variety of flaws in design and analysis and should be interpreted with caution.

• Height Change in Children
  o Evidence on DEX vs MPH or MPH IR alone is inconsistent. No conclusions can be drawn:
    o Limited evidence suggests that height changes resulting from atomoxetine are similar to those reported with MPH IR, and are also transient.

• Weight in children
  o DEX vs MPH: Results from comparative observational studies suggest that DEX is associated with significantly greater suppression of weight gain than MPH in the first 1-2 years. However, the difference between DEX and MPH appears to resolve by the second year and the difference found in years 1-2 may have been exaggerated by higher relative DEX dosages. Ultimately, these data should be interpreted with caution, due to methodological flaws in the measurement of weight.
    o The remaining comparative and noncomparative observational studies do not support a definite relationship between MPH and suppression of weight gain.
    o Limited evidence suggests that weight changes resulting from atomoxetine are similar to those reported with MPH IR, and are also transient.

• There is no comparative evidence on other long-term safety outcomes, including tics, seizures, cardiovascular adverse events, injury frequency, and hepatotoxicity.
• Warnings and Removal from Market:
  o Adderall XR® removed from market in Canada due to reports of sudden death in children.
  o Atomoxetine: reports of severe hepatotoxicity led to additional warnings in product label
  o Pemoline: Reports of acute liver failure led to boxed warning in product label – not to be used as first line therapy

Adolescents

Efficacy and tolerability
• Adolescents were studied in a small number of short-term trials that involved MPH IR, MPH OROS (Concerta®) or pemoline. No studies of nonstimulants reported effects in adolescents as a separate group, studies of atomoxetine included adolescents and are discussed above.
• MPH OROS vs MPH IR
  o A single, very small, single blinded, study showed MPH OROS superior to MPH IR on some measures of simulated driving skills, during tests administered in the late evening or nighttime. No difference was found during other test times.
• Indirect Evidence: Stimulants
  o Placebo-controlled trials of MPH IR and pemoline do not provide indirect evidence of comparative efficacy or tolerability due to heterogeneity in outcome reporting.
  o MPH IR generally was superior to placebo in improving core ADHD symptoms, but was associated with more frequent reports of appetite and sleep disturbances.
  o Pemoline was more efficacious than placebo in reducing ADHD symptoms in one study; but had insignificant effects in another study involving adolescents with comorbid substance use disorder. Pemoline was frequently associated with insomnia in both trials.
• Functional outcomes: observational studies
  o Observational studies of MPH IR that report functional outcomes found mixed results. In an uncontrolled study of young adult males who had taken MPH as children and teens (mean age at discontinuation of MPH 17 years) fewer suicide attempts were associated with higher dose of MPH. Emancipated living situation and level of relationship commitment was associated with response to MPH. Early response to MPH was negatively associated with high school graduation, however.
  o Another uncontrolled follow-up of MPH IR responders reported “improved grades” after 6 – 14 months. Methodological limitations of these studies severely limit the interpretation these findings.

Long-Term Safety
• We found no evidence on long-term safety of drugs used to treat ADHD in adolescents
Adults

Efficacy and tolerability

- Pharmacological treatment of ADHD in adults has not been widely studied.
- Adderall, DEX IR, and MPH IR lead to response in 57-70.4% of participants in placebo-controlled trials.
- There is no evidence that any one stimulant is more effective than any other.
- Pemoline has been associated with liver toxicity and is not recommended for first-line therapy. Response rates were 38.5-50% in placebo-controlled trials.
- Other pharmacologic treatments have emerged that offer viable alternatives to stimulant therapy and response rates are as follows:
  - Atomoxetine: 52.4%
  - Bupropion SR: 64-76%
  - Bupropion XL: 53%
  - Guanfacine: Unknown, but positive effects on ADHD-RS scores were similar to DEX
  - Modafinil: 48%
- Evidence regarding treatment effects on quality of life and other ADHD-related symptoms (depressed mood, anxiety, and cognition) in adults is not compelling.
- Short-term, randomized controlled trials do not provide clear evidence that any one stimulant is more tolerable than another or that nonstimulants offer an advantageous tolerability profile over stimulants.
- No conclusions about comparative efficacy or safety in subgroups of adults can be made.

Long-term safety

- We found one follow-up study of participants in an MPH IR RCT that reported a single attempted suicide among 8 men (12.5%).

Subgroups

- Race / Ethnicity
  - Only ½ of studies reported race or ethnicity data; Studies were primarily conducted in white populations.
  - 2 placebo-controlled studies in 100% non-White groups:
    - MPH IR in African American boys
      - 75% of subscale measures showed improvement
      - This rate is similar to response rates reported in other trials
      - Linear increases in diastolic blood pressure noted
    - Clonidine in 10 children in India with comorbid mental retardation
      - Improvements compared to baseline on 90% of measures
- Gender
  - No difference in efficacy can be found between boys and girls.
- Commonly occurring comorbidities:
  - ½ of studies reported, but none stratified analyses
  - Prevalence in studies (AAP estimated prevalence)
    - oppositional defiant disorder: 19-66.7% (35.2)
    - conduct disorder 9-38.5% (25.7)
    - anxiety 1.4-42% (25.8)
    - depression 0.7-6.6% (18.2)
Adults: Atomoxetine and pemoline: Subgroup analyses of placebo-controlled trials suggested that presence or absence of co-occurring “Psychiatric illness” did not alter treatment effects.

- **Tics**
  - Placebo-controlled studies of MPH IR do not consistently support a relationship to increased tic severity or frequency. A few measures improved or worsened, but global measures and total scores do not show a difference.
  - In direct comparison, clonidine improved tic symptom scores compared to MPH IR, but no difference was found on the primary ADHD symptom outcome.
  - Limited evidence does not show a difference between clonidine and placebo in tic severity or frequency. Equally limited evidence suggests guanfacine reduces tic severity compared to placebo.

- **Mental Retardation / Developmental Delay**
  - In children with mental retardation, evidence indicates that MPH IR is beneficial on most ADHD outcomes compared to placebo.
  - Adverse events were common, with staring and social withdrawal occurring more often with MPH IR than placebo.
  - Limited evidence also supports a benefit for clonidine.
  - Drowsiness and blood pressure lowering were seen more commonly with clonidine than placebo. No comparative evidence is available.

- **Autism or Epilepsy**
  - Very limited evidence is available on the use of MPH IR in children with autism or epilepsy. However, this evidence suggests that MPH IR is beneficial on most ADHD outcomes compared to placebo. This evidence should be interpreted with caution.

- **Substance Abuse**
  - Very limited evidence suggests that pemoline is beneficial for some outcome measures of ADHD in substance abusing teens, but it did not affect substance use or conduct disorder symptoms. This evidence should be interpreted with caution.
  - Very limited evidence suggests that response rates to MPH IR in Adults: with co-occurring cocaine dependence was similar to rates found in other trials.

- **No conclusions about comparative effectiveness or safety based on age, gender, race/ethnicity or comorbidities can be made from this body of evidence.**
Detailed Assessment

Key Question 1: What is the comparative efficacy or effectiveness of different pharmacologic treatments for attention deficit disorders?

Young Children (Preschool Age; 3-5 years)

Evidence on the effectiveness of pharmacotherapy for ADHD in young children is seriously lacking (Evidence Tables 1 and 2). We did not find any effectiveness trials, long-term observational studies assessing functional outcomes or efficacy trials comparing drugs in young children with ADHD.

The evidence of any short-term benefit of stimulants in this age group comes from 48-52 placebo-controlled trials. However, only 1 trial was fair quality and used an assessment tool with good validity (CPRS-R; learning, conduct and hyperactivity indices only).50 In this study, both the high dose (0.5 mg/kg twice daily) and the low dose (0.3 mg/kg twice daily) resulted in lower scores than while on placebo at the end of 7 to 10 days of treatment. The high dose resulted in better final scores than the low dose on only the learning component of the CPRS-R with the low dose resulting in a mean of 8 points (10%) lower, and the high dose a mean of 14 points (18%) lower than the score while on placebo. The clinical importance of these differences is not known, and baseline scores are not reported or accounted for. Based on parental report, medication did not result in better compliance with tasks compared to placebo, although reports of time on task were better with the higher dose (mean 52 seconds longer compared to placebo). The DSM-III criteria were used to diagnose ADHD. ADHD subtypes or ethnicity were not identified in this study. MPH was associated with higher rates and greater severity of adverse events than placebo, significantly more in the higher dose group. Rates of specific adverse events were not reported. We found no evidence on long-term safety of drugs used to treat ADHD in young children.

Children (Elementary School Age; 6-12 years)

Generalizability Issues

Studies of elementary school age children with ADHD were characterized by under-reporting of baseline subtype classifications, race or ethnicity, co-occurring disorders, and illness severity. This gap in the literature limits the generalizability of the findings to target populations. Only one-quarter of all studies of school-aged children reported ADHD subtype prevalence rates. The mixed subtype was most common, occurring in 58-100% of participants across most study populations. The inattentive subtype was generally observed less frequently (prevalence rate range: 9-40%) and the hyperactive subtype was relatively rare (prevalence rate range: 1-8%). Only one-half of all studies of elementary school-aged children reported race or ethnicity among the baseline characteristics. The racial/ethnic make-up of the majority of these study populations was consistent with the current U.S. Census Bureau Estimates (white=80.4%, black=12.8%, Asian=4.2%, and of Hispanic/Latino origin=14.1%).53 However, the prevalence of ADHD among ethnic groups may not correlate with these data.

Just over half of studies reported prevalence rates of co-occurring disorders, including oppositional defiant disorder (19-66.7%), conduct disorder (9-38.5%), anxiety (1.4-42%) and depression (0.7-6.6%). With the exception of depression, the ranges of comorbidities reported in these trials encompass the American Academy of Pediatrics estimates on prevalence of common comorbidities: Oppositional defiant disorder=35.2 (27.2, 43.8), conduct disorder=25.7 (12.8, 41.3), anxiety disorder=25.8 (17.6, 35.3) and depressive disorder=11.1, 26.6).54 Illness severity
was not presented as a baseline characteristic in most studies, and comparisons across studies based on scales used to assess symptoms are hampered by variation in scale choice and method of reporting. Diagnostic processes also varied across studies. Seventy-two percent of studies used either the DSM III, DSM III-R, or DSM IV criteria to diagnose ADHD, however many used additional criteria and the clinical comparability of patients enrolled is not clear.

**Stimulants**

**METHYLPHENIDATE**  
**Comparison of Immediate Release and Sustained Release Formulations**

We included seven trials of MPH IR versus SR formulations.\textsuperscript{25, 55-61} Of these, four were poor quality due to either inadequate or undescribed methods of randomization and allocation concealment, combined with lack of description of an intention to treat (ITT) analysis, lack of information on eligibility criteria, attrition, or post-randomization exclusions (Evidence Table 3).\textsuperscript{25, 55, 56, 61} The remaining 3 studies compared MPH OROS (Concerta®) and MPH SR (Ritalin SR®) to MPH IR.\textsuperscript{57, 59, 60} No trials comparing the other extended release formulations of MPH (Ritalin LA®, Methylin ER®, Metadate ER®, or Metadate CD®) to MPH IR were found.

**MPH IR versus MPH OROS (Concerta®)**

Including the largest RCT, with 312 children,\textsuperscript{59} 2 studies of MPH IR versus MPH OROS did not show an overall difference in outcomes\textsuperscript{57-59}. The RCT found no significant differences between the formulations on the primary outcome measure (IOWA Conners scale) or on 11 secondary measures. While not reaching the level of ‘poor’ quality, some concerns about the study were identified. Although the randomization was stratified, the analyses were not and all data from patients enrolled at one study site were excluded from the analysis due to ‘irregularities in the data collection’. There were small differences between groups at baseline, including larger proportions of patients with comorbidities and a diagnosis of inattentive ADHD in the MPH IR group, and larger proportions of patients with a diagnosis of combined ADHD and who were taking MPH IR prior to enrollment in the MPH OROS group. The baseline inattention/overactivity score on the IOWA Conners parents scale was significantly higher in the MPH OROS group than the MPH IR group.

A much smaller crossover trial (70 children), 7 days long, found MPH OROS to have lower scores on the Abbreviated Conners Parents scale (total), and on the inattention/overactivity item (out of 16 items), however no differences were found based on assessments made by teachers and counselors.\textsuperscript{57, 58}
**MPH IR versus MPH SR (Ritalin SR®)**

In a small 2-week RCT (34 children) of MPH IR versus MPH SR found mixed results. The outcome measures included questionnaires (not validated) completed by a physician, a teacher and a parent. The teacher questionnaires indicated significant differences in final total score and the “Conduct Problem” score favoring MPH IR. Parent questionnaires indicated a significant difference favoring MPH SR on the “Conduct Problem” item final score, and the physician scores showed no difference.

**Other Measures of Comparative Effectiveness of IR vs SR formulations**

Clinical trials of extended release versus immediate release formulations were too short to demonstrate differences in long-term health outcomes. However, the intermediate outcome measure of persistence (the proportion of patients continuing to take or refill prescriptions for a medication after some longer period of time) is thought to be a good proxy for extension of benefits seen in the short-term, or if none were found, evidence of a difference in long-term, real-life settings. Persistence is an intermediate outcome with unknown validity because direct evidence of a relationship between persistence rates and long term health outcomes with ADHD drugs is lacking.

Two observational database studies reported persistence outcomes for 12-month periods following index prescriptions of MPH IR and ER formulations. MPH ER formulations were associated with better persistence outcomes than MPH IR in both studies regardless of measurement methods. The findings of these studies should be interpreted with caution, however, until confirmed by a randomized controlled trial that would serve to rule out potential sources of bias, including between-group baseline differences in unmeasured clinical characteristics, physicians’ prescribing preferences, and differences in reasons for discontinuation (e.g., change in insurance benefit, use of promotional samples).

Data were derived from the Integrated Health Care Information Services (IHCIS) National Managed Care Benchmark Database in one study that reported the proportion of 1,775 patients that persisted with their index prescription for 12 months with no discontinuations exceeding 14 days. MPH OROS was associated with greater persistence rates than MPH IR (12% vs 1%, p<0.0001). There is uncertainty about how well this study population represents patients in actual practice as ethnicity and comorbidity characteristics were not reported.

California Medicaid claims files from a 3-year period were examined in the second study to identify youth prescribed MPH (n=11,537). This study population involved a lower than average proportion of white patients (45.3%) and higher proportions of Hispanic patients (26.1%). Total mean duration (days) of treatment without any 30-day gaps was greater for patients taking ER formulations (combined group of MPH OROS = 83%, MPH ER = 8.7%, MPH SODAS = 8.3%) than for those taking MPH IR (140.3 vs 103.4; survival time ratio (STR) 1.37, 95% CI 1.32-1.42). Subgroup analysis results suggest that persistence duration was greatest for MPH OROS (147.2 days, 95% CI 142.6-151.7 days) compared to MPH SODAS (113 days; 95% CI 100.9-125.1 days) or MPH CD (101.1 days, 95% CI 91.2-111.0 days). Together, ER formulations extended persistence duration regardless of ethnicity.
Comparisons of SR Formulations

MPH OROS (Concerta®) vs MPH SODAS (Ritalin LA®)

A small 1-week crossover study of MPH SODAS 20mg versus MPH OROS 18mg and 36mg found MPH SODAS superior on the attention or deportment subscores of the Swanson, Kotin, Agler, M-Flynn and Pelham (SKAMP) scale depending on the time-point and dose comparison. Secondary outcome assessment also found MPH SODAS superior on one measure (proportion correct on math test). These limited differences are mitigated by concerns over the assessment tool (SKAMP) sensitivity, use of a simulated classroom, involvement of study sponsor in authorship, and differences in groups at baseline.

No direct comparisons of other extended release formulations of methylphenidate or other ADHD drugs were found.

Methylphenidate ER (Metadate®) vs Placebo

A 3-week trial of Metadate® versus placebo enrolled 314 children out of 507 screened. Twenty four percent of those excluded at screening were because they responded to placebo during a 1-week washout period. This biases the study population towards the Metadate® arm, reducing the applicability of the results. The mean change in the primary outcome measure, the teachers CGI ratings combined, in the morning and afternoon, were significantly lower (better) in the Metadate® group. Secondary measures also favored Metadate®.

Immediate Release Formulations

Dextroamphetamine versus Methylphenidate

Nine fair quality studies (reported in 11 publications) of DEX versus MPH IR found two poor quality studies, and one poor-quality sub-group analysis were found. All nine fair quality studies were randomized, blinded crossover trials. Table 3 summarizes the study characteristics.
The two largest studies,\textsuperscript{32,65} which used clear criteria for diagnosis, enrolled children with ADHD in order to test the hypothesis that some adverse events associated with stimulants are actually characteristics of ADHD and would be improved by drug treatment in one study,\textsuperscript{65} and to test the differences between child and parent assessment of therapy in the other.\textsuperscript{32} Neither study provides details on the efficacy results, other than summary statements that there were no differences between the two drugs based on children’s self-assessment\textsuperscript{32} and based on parent and teacher ratings.\textsuperscript{65} These 2 studies had similar populations, primarily children with the Mixed subtype (82%), however co-morbidities and ethnicity are not reported.

Of the 7 small studies (n = 12 to 48), only one found a difference between the drugs.\textsuperscript{45} This study assessed attention to task and deviant behavior in the usual classroom settings using a modified version of the Werry-Quay Direct Observational System.\textsuperscript{45} The text of the paper reports that in a post-hoc analysis, DEX was the most effective drug \textit{in instances where a positive effect was seen}. Because this study did not use a standardized tool for diagnosis, and ADHD subtypes, co-morbidities or ethnicity are not reported, it must be assumed that significant heterogeneity in the population may have lead to the discordant results.
Response Rates

Very few studies attempted to make a comparison of the rate of response (defined a priori) between 2 drugs. Table 4 shows the studies that did. Overall, no differences in response rates, as defined below, were found between the comparisons of MPH OROS, DEX IR, Amphetamine mixture and clonidine to MPH IR. Additionally, the majority of these response rates are lower than those reported and quoted from placebo controlled trials (rates of approximately 75%).

<table>
<thead>
<tr>
<th>Interventions</th>
<th>Response Rate Definition</th>
<th>Response rates (% pts)</th>
</tr>
</thead>
<tbody>
<tr>
<td>MPH OROS vs MPH IR</td>
<td>Parent/teacher ratings of Global Effectiveness as &quot;Good&quot; or &quot;Excellent&quot;</td>
<td>Parent: 67.2 vs 64.7 Teacher: 67.2 vs 57.4</td>
</tr>
<tr>
<td>Pelham 2001</td>
<td>MPH OROS</td>
<td>67.2 vs 64.7</td>
</tr>
<tr>
<td>Crossover N=70</td>
<td>MPH IR</td>
<td>67.2 vs 57.4</td>
</tr>
<tr>
<td>Wolraich 2001</td>
<td>MPH OROS</td>
<td>CI rated as &quot;much&quot; or &quot;very much&quot; improved</td>
</tr>
<tr>
<td>Parallel N = 192</td>
<td>MPH IR</td>
<td>46.2 vs 47.2</td>
</tr>
<tr>
<td>DEX IR vs MPH IR</td>
<td>Parent/teacher ratings of Global Effectiveness as &quot;Good&quot; or &quot;Excellent&quot;</td>
<td>Parent: 54 vs 46.5 Teacher: 42.9 vs 46.9</td>
</tr>
<tr>
<td>Efron 1998</td>
<td>DEX IR</td>
<td>62.4 vs 73.5</td>
</tr>
<tr>
<td>Crossover N=102</td>
<td>MPH IR</td>
<td>62.4 vs 73.5</td>
</tr>
<tr>
<td>Efron 1997</td>
<td>DEX IR</td>
<td>68.8 vs 72</td>
</tr>
<tr>
<td>Crossover N=125</td>
<td>MPH IR</td>
<td>68.8 vs 72</td>
</tr>
<tr>
<td>Sharp 1999</td>
<td>DEX IR</td>
<td>85 vs 83</td>
</tr>
<tr>
<td>Crossover N=42</td>
<td>MPH IR</td>
<td>85 vs 83</td>
</tr>
<tr>
<td>Amphetamine mixture (Adderall) vs MPH IR</td>
<td>CGI improvement score of 1 or 2: &quot;very much improved&quot; or &quot;much improved&quot;</td>
<td>90 vs 65; p=0.12</td>
</tr>
<tr>
<td>Pliszka, 2000</td>
<td>Adderall:</td>
<td>90 vs 65; p=0.12</td>
</tr>
<tr>
<td>Parallel N = 40</td>
<td>MPH IR</td>
<td>90 vs 65; p=0.12</td>
</tr>
<tr>
<td>Clonidine vs MPH IR</td>
<td>Clinical responders: behavior improved at home/or at school</td>
<td>47 vs 53</td>
</tr>
<tr>
<td>van der Meere 1999</td>
<td>Clonidine</td>
<td>47 vs 53</td>
</tr>
<tr>
<td>Parallel N=53</td>
<td>MPH IR</td>
<td>47 vs 53</td>
</tr>
</tbody>
</table>

Functional Outcomes:

We found extremely limited information on functional capacity outcomes from the clinical trials. Therefore, we included observational studies of >/= 6 months duration that reported outcomes that reflect functional capacity, for example academic achievement in terms of progression through grades, suicide attempts, police contacts, etc. We found only 2 studies (Evidence Tables 13 and 14).

MPH IR vs MPH OROS (Concerta®)

IHCIS managed care claims’ data (described above) suggest that MPH OROS was associated with fewer outpatient visits/hospitalization for accidents/injury than MPH IR over a 12-month follow-up period (Odds ratio 0.58, 95% CI 0.353 to 0.945). The study population
was 75% male, with a mean age of 9.7 years; however, no other information regarding ADHD subtypes, comorbidities, or race/ethnicity were provided.

MPH IR

In a 4-year follow-up study of 62 children treated with MPH, the effect of duration of treatment on academic performance was assessed. The duration of treatment was divided into < 6 months, 6 months to 2 years, 2 – 3 years, and 3 – 4 years, and those currently taking stimulants at follow-up. No differences were found between the groups on academic achievement as measured by teachers, the proportion repeating grades, in special education classes or being tutored. Although the proportion of children repeating grades was lowest in the group continuing to take MPH (8% vs 46%, 50%, 36%, 31%) this difference was not statistically significant – possibly because of the small numbers of boys per group (10 to 14). Due to methodological limitations, this study provides no comparative information.

Maintenance of Short-Term Symptom Response Effects

MPH or DEX versus Placebo or Non-drug Therapy

All of the trials reported above are very short-term trials (range 1 to 9 weeks). Because of this serious limitation, this evidence does not provide information on the long-term benefits of these drugs in treating ADHD. To provide further evidence on duration of effect, and longer-term outcomes, placebo- or non-drug therapy controlled trials of ADHD drugs with duration \( \geq 6 \) months are reported here (Evidence Tables 7 and 8).

We found 3 placebo-controlled trials of at least 6 months duration, 1 with DEX IR and 2 with MPH IR, and 3 trials that randomized children to stimulant medication or non-drug therapy for 12 to 14 months. Two studies were poor quality due to serious flaws that represent significant potential for bias. The placebo controlled trial of DEX did not report any baseline characteristics of the two groups, and did not conduct an ITT analysis while the numbers and reasons for withdrawal are also not reported. A trial of MPH IR, cognitive training or both (n=30) omitted important information about basic information on study design and outcomes (e.g. randomization, baseline characteristics, blinding, and loss to follow up).

Overall, the MPH IR studies provide a mixed picture of the consistency of efficacy of MPH over 6 months to 2 years. The only study reporting that the short-term effects were maintained over the follow-up period was the Multimodal Treatment Study of Children with Attention Deficit/Hyperactivity Disorder (MTA) study.

The Multimodal Treatment Study of Children with Attention Deficit/Hyperactivity Disorder (MTA) was a relatively large study (n = 579) funded by the NIMH assessing medication management, behavioral treatments, standard community care, and combined medication management and behavioral treatments over a 14-month period. Outcomes are available for 540 children that were followed an additional 10 months subsequent to trial discontinuation. Medication management could involve any stimulant medication, but started with MPH titration. At study end, 73% of those in the medication management groups were on MPH, and 10% on DEX, with small numbers of patients taking no medication, pemoline, imipramine, bupropion or haloperidol and 6% refusing to be in the medication arm assigned. All participants met DSM-IV criteria for ADHD combined type, had a mean age of 8.5 years and 80% were males. The sample population was ethnically diverse, with White (61%), African American (20%), and Hispanic (8%) representation. Comorbidities included anxiety disorder (33.5%), conduct disorder (14.3%), oppositional-defiant disorder (39.9%),
affective disorder (3.8%), tic disorder (10.9%), mania/hypomania (2.2%) and other (e.g., bulimia, enuresis) (0.2%).

Medication management alone resulted in better scores compared to behavioral therapy for the symptoms of inattention (both parents and teachers) and hyperactive-impulsive symptoms (parent ratings). Medication alone resulted in better scores on all ADHD symptoms than community care, except as measured by a classroom observer. Aggression-ODD symptoms scores were better with medication alone compared to community care in teacher ratings only. Combined therapy (medication and behavioral therapy) was not different to medication alone on any scale. Important to this review of ADHD medications, the effect of medication management was maintained over the 14 month period. This study was a pragmatic trial in that the treatments were given openly (after blinded titration in the 2 drug treatment arms), and participants could refuse the assigned arm, or add or change treatments. In the community care arm, for example, 68% were taking ADHD medications although the mean dose and number of doses per day of MPH was lower in the community care arm than the medication arms. However, the outcome measures were not effectiveness outcomes, so the trial must still be viewed as an efficacy trial that indicates that with careful monitoring of dose and drug regimen, ADHD stimulant medications can reduce symptoms of ADHD over a 14-month period.

Families were contacted 10 months after the end of the 14-month study (24 months post-randomization) to assess longer-term persistence of treatment effects. A total of 540 of the originally randomized 579 (93%) participated and 10 months after study end; 72% in the medication management alone group, 70% in the combined therapy group, 62% in the community care group were taking medication for ADHD. At 24 months post-randomization, medication alone resulted in better scores on ADHD and ODD symptoms than behavioral therapy and community care. Despite this, analyses of combined outcomes from the medication management alone and combined therapy groups compared to those of the behavioral therapy and community care groups suggest a reduction in the improvement magnitude by half from the 14-month to 24-month timepoints; effect size changes for ADHD symptoms=0.60 vs 0.30 and ODD symptoms=0.39 vs 0.21.

The other earlier trials reported a dissipation of effect over time (Table 5). Although some of these studies do not report mean doses, of those that do the doses used in the MTA study were higher.

### Table 5. Maintenance of MPH IR short-term effects

<table>
<thead>
<tr>
<th>Study</th>
<th>Treatments Duration</th>
<th>Sample Size</th>
<th>Mean age (yrs)</th>
<th>% Male</th>
<th>Results</th>
</tr>
</thead>
<tbody>
<tr>
<td>Kupietz 1998</td>
<td>MPH IR 0.3, 0.5 or 0.7 mg/kg Placebo x 27 weeks</td>
<td>N=47</td>
<td>9.7</td>
<td>9.7 % Male NR</td>
<td>Mean CTRS total ratings worse at Week 27 than Weeks 2 or 14</td>
</tr>
<tr>
<td>Ialongo 1993</td>
<td>MPH IR 0.4 or 0.8 mg/kg Multimodal treatment MPH 0.4 mg/kg + Multimodal Placebo Trial=3 months; 9 months follow-up</td>
<td>N=96</td>
<td>8.27 years</td>
<td>77.4%</td>
<td>Short-term gains deteriorated at 9 months</td>
</tr>
<tr>
<td>Firestone 1986</td>
<td>MPH IR 22 mg Parent training Both x 2 years</td>
<td>N=73</td>
<td>Mean age NR</td>
<td>Gender NR</td>
<td>Conners Hyperactivity Index Scores worsened at 1 (N=52) and 2 (N=30) years</td>
</tr>
</tbody>
</table>
Remission Rates: MPH IR

Three studies assessed the effects of withdrawing MPH IR after periods of treatment. Two of these were poor quality, but the third study included a group of 21 boys who had been treated with MPH for a mean of 1.75 years and randomized to 3 weeks of placebo or MPH. Using the CTRS, this study found that on the Subscale items of hyperactivity and defiance the scores during the placebo period were significantly worse than during the MPH period. No baseline assessments were presented, and the analyses are based on scores at week 3 of each condition only so there is no information about the effectiveness of their pre-existing MPH regimen at baseline. In addition, the effect of order of drug/placebo was not analyzed in this crossover study, so the results must be interpreted with caution.

OTHER STIMULANTS

Amphetamine mixture versus Methylphenidate Immediate Release

Three small, fair quality studies of Amphetamine mixture versus MPH IR were found. One was a parallel group RCT while the other two were randomized cross-over trials. Two additional studies were rated poor quality due to no description of randomization or concealment of randomization code, no ITT analysis, and high discontinuation rates or no randomization (clinician selected drug) and no blinding of patients or outcome assessors.

The parallel group RCT enrolled 58 children with ADHD and randomized to 3 weeks of Amphetamine mixture, MPH IR or placebo. The mean doses at the end of study were Amphetamine mixture 12.5 mg/day and MPH IR 25.2 mg/day (divided into morning +/- noon doses for both drugs). No differences were found in the mean IOWA CTRS scores (Inattention/Overactivity and Aggression/Defiance subscales) rated by teachers 4 mornings and afternoons a week, but amphetamine mixture was significantly better on both subscales when morning and afternoon scores were combined. No differences were found in parent ratings. The mean CGI-Improvement score (rated by a blinded psychiatrist) was also significantly lower (better) in the amphetamine mixture group than the MPH IR (final score 1.6 vs 2.35, p<0.05), but the difference in the proportions of responders (90% vs 65%, respectively) did not reach statistical significance. No differences were found on the Conners Global Index or final weight.

The two crossover studies were conducted in the same manner by the same authors, conducted in a summer treatment program. These short-term studies (6 – 8 weeks) enrolled 21 and 25 children, with a higher prevalence of comorbid oppositional defiant disorder (67% and 52%) than the general population of children with ADHD. The first study found amphetamine mixture to be superior to MPH IR given once daily, while few or no differences were found when comparing to MPH IR given twice daily, based on counselor and teacher ratings. Parent ratings of after school behavior indicated that the addition of a third 0.3mg/kg dose of MPH IR, or the amphetamine mixture 0.3 mg/kg once daily dose lead to the best results based on combinations of parent ratings and child task completion. The results of the second study indicate that on a few measures the low dose (10mg twice daily) of MPH IR was not as effective as the higher dose (17.5 mg twice daily) or either dose of amphetamine mixture (7.5 or 12.5 mg twice daily). Measures where this difference was seen were interruption, conduct problems, negative verbalizations, the daily report card score, and counselor ratings of oppositional defiant scores. No difference in response was seen between the two doses of amphetamine mixture and the higher dose of MPH IR.
Amphetamine Mixture versus Dextroamphetamine

The evidence is limited to a single poor quality study of dextroamphetamine IR versus dextroamphetamine SR versus Amphetamine mixture versus placebo. No conclusions can be drawn.

Pemoline versus MPH or DEX

The evidence for the comparative efficacy of pemoline is limited to 3 poor quality trials. Two trials of pemoline versus MPH IR were rated poor quality due to multiple serious flaws (see Evidence Table 4). A study of pemoline versus dextroamphetamine was rated poor quality for similar reasons. No conclusions can be drawn about the comparative efficacy of pemoline.

Dexmethylphenidate (d-MPH)

While 2 placebo-controlled studies of d-MPH are referred to in the most recent FDA Medical Review (http://www.fda.gov/cder/foi/nda/2001/21-278_Focalin_medr_P1.pdf), neither of these were found in the published literature, and no dossier was submitted by the manufacturer. A small study of the effects of withdrawing d-MPH after a 6-week titration period was poor quality. No conclusions can be drawn about the comparative efficacy of d-MPH.

Modafinil

A small study of modafinil enrolled 24 children with ADHD. The study duration was variable, with mean durations of 5 and 6 weeks (placebo and modafinil, respectively). In this study, less than 1/3 had oppositional defiant disorder, or conduct disorder (27% combined), and the ADHD subtype was primarily Mixed (73%). Two children (8%) in the modafinil group were excluded from the analysis because they did not have post-randomization assessments. Modafinil was found better than placebo in improving scale scores, and parent assessments.

Non-stimulants

ATOMOXETINE

Atomoxetine versus Methylphenidate

Atomoxetine, the first nonstimulant introduced specifically for ADHD was compared to MPH IR in 3 RCTs. However, 2 of these studies were really comparisons to placebo, with only few patients enrolled in the MPH arms. Therefore, these are considered placebo-controlled trials, below. The single study comparing atomoxetine and MPH IR found no differences between the drugs based on changes in the ADHD-RS, the CPRS-R hyperactivity item, and the CGI-S. Concerns over the study quality indicating potential bias suggest caution in interpreting these findings (see Evidence Table 4).

Atomoxetine versus MPH OROS

From the review done for NICE, we also know of a trial comparing MPH OROS with atomoxetine, which has not yet been published. Study details submitted to NICE for consideration were not submitted by sponsor (McNeil Pharmaceuticals) for this review. Although much of the data submitted was removed from the NICE report due to commercial considerations, the review does state that the MPH OROS resulted in better scores on the ADHD-RS and the CGI scale.
Atomoxetine versus Placebo

Five placebo controlled studies of atomoxetine in children and adolescents with ADHD found atomoxetine to be superior based on ADHD-RS as the primary outcome measure and various scales as secondary measures. The mean change on ADHD-RS in these 6 to 9 week studies ranged from –12.8 to -16.7 with atomoxetine compared to –5 to –7.0 for placebo. A study of once daily dosing reported response rates (defined as >/= 25% reduction in ADHD-RS score), in the atomoxetine group of 59.5% versus 31.3% in the placebo group (p<0.001). Remission rates (defined as an endpoint CGI-S score of 1 or 2) were 28.6% and 9.6%, respectively (p=0.003). All 5 studies were funded and co-authored by representatives of the manufacturer of atomoxetine, and 4 were part of the NDA submitted to the FDA. All used the DSM IV criteria, however the proportions of ADHD subtypes varied, for example 52 to 79% of enrolled children had the Mixed subtype. More concerning is the variation in the proportions of children with each subtype per assigned group. Proportions of children with co-morbidities also varied across the studies (e.g. 18 to 45% with oppositional defiant disorder). No analyses based on subtype or co-morbidities were done.

A study of the effect of continuing atomoxetine (or switching to placebo) in children who had been classified as responders in an open label study enrolled 416 children for 9 months. Relapse was defined as return to 90% of baseline ADHD-RS score and CGI-S score increase of at least 2 points. The primary outcome measure was the number of days to relapse. The mean number of days to relapse was 218 with atomoxetine and 146 with placebo (p<0.001). Similarly, fewer patients on atomoxetine relapsed than on placebo (22% versus 38%, p<0.002).

BUPROPION

Bupropion IR versus Methylphenidate

The antidepressant bupropion was compared to MPH IR in a randomized crossover trial of 18 children with ADHD. This study was somewhat unusual in that one-third of children had developmental learning disorders, and lower rates of comorbidity for conduct disorder or oppositional defiant disorder (13% each). Overall, the findings of the study do not demonstrate a significant difference between the drugs. An analysis of the effect of the order of randomization revealed that the drug that was received first resulted in the larger effect (p= 0.03). Because the results are not reported separately by order of randomization, these findings suggest that the results combined analysis may not be valid.

Bupropion IR versus Placebo

Two trials of bupropion versus placebo were found. Both had authors from the manufacturer of bupropion, and one reported funding from this manufacturer. Both trials had concerning flaws, including lack of reporting on method of randomization, concealment of random orders, blinding of outcome assessors, and attrition rates, and lack of clarity on numbers enrolled. The better of these 2 studies found inconsistent results, depending on the rater (parent or child) and the scale or Subscale.

CLONIDINE

Clonidine versus Methylphenidate or Placebo

Three parallel group RCTs of clonidine versus MPH IR were found. A study of children with Tourette’s disorder and ADHD enrolled 136 children and assigned them to MPH IR, clonidine, both drugs, or placebo for 8 weeks. Mean doses at the end of study were 0.25 mg clonidine, 26 mg MPH IR daily. All analyses made comparisons of
each drug group to placebo; although it is stated that there was no difference between the MPH IR and clonidine groups on the primary outcome measure of the Conner-ASQ-Teacher scale. Clonidine was significantly better than placebo on more Tourette’s outcome measures than MPH IR, and MPH IR was significantly better than placebo on more ADHD outcome measures than clonidine.

**Table 6. Summary of Differences in Results of ADHD/Tourette’s Disorder Study**

<table>
<thead>
<tr>
<th></th>
<th>Clonidine vs Placebo</th>
<th>MPH IR vs Placebo</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>ADHD SYMPTOMS</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>IOWA Total score</td>
<td>No Difference</td>
<td>Drug Found Significantly Better</td>
</tr>
<tr>
<td>IOWA OD</td>
<td>No Difference</td>
<td>No Difference</td>
</tr>
<tr>
<td>Classroom Observation: On Task</td>
<td>No Difference</td>
<td>Drug Found Significantly Better</td>
</tr>
<tr>
<td>Conners CPT Attentiveness</td>
<td>No Difference</td>
<td>No Difference</td>
</tr>
<tr>
<td>Conners CPT Risk Taking</td>
<td>No Difference</td>
<td>No Difference</td>
</tr>
<tr>
<td><strong>Tourette’s SYMPTOMS</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>YGTSS Motor</td>
<td>Drug Found Significantly Better</td>
<td>No Difference</td>
</tr>
<tr>
<td>YGTSS Vocal</td>
<td>Drug Found Significantly Better</td>
<td>No Difference</td>
</tr>
<tr>
<td>GTRS Teacher</td>
<td>No Difference</td>
<td>No Difference</td>
</tr>
</tbody>
</table>

Two small RCTs of clonidine versus MPH IR measured outcomes using scales or tests that have either been show to have low validity (e.g., Home/School Situations Questionnaire (HSQ and SSQ), Gordon Diagnostic System), or which validity could not be verified (e.g., Disruptive Behavior Scale, Grooved Pegboard). A small randomized placebo-controlled cross-over trial was rated poor quality while another did not report clear eligibility criteria but it appears that the children had to have a diagnosis of ADHD and Tourette’s disorder based on DSM-IIIR. Outcomes were assessed using primarily Child Behavior Checklists for parents and teachers, and linear analogue scales of parent assessments of hyperactivity and tics, as compared to any time in the past. Based on the linear analogue assessments, clonidine was not significantly different to placebo. The assessment of tics, based on 4 scales, did not show a difference between placebo and clonidine. Multiple subgroup analyses of the checklists resulted in clonidine being superior to placebo on some subscale items.

**GUANFACINE**

A small study of 24 children with ADHD, all of the mixed type, and a tic disorder studied the effects of guanfacine versus placebo for 8 weeks. Slightly more than half of enrolled children had Tourette’s disorder (58.8%), and 35% had chronic motor tic disorder. Teachers and investigators rated guanfacine superior to placebo, while parent ratings did not.

**RISPERIDONE**

The only evidence of risperidone’s effects on ADHD in children comes from post-hoc subgroup analyses of studies designed for treatment of children with subaverage IQ’s and disruptive behavior disorder. Results from a meta-analysis of subgroups from 2 six-week placebo-controlled trials that involved 155 children with comorbid ADHD suggest that risperidone was associated with reductions in Nisonger Child Behavior Rating Form (NCBRF) Hyperactivity subscale scores, regardless of concomitant stimulant use. Results from another post-hoc analysis suggest that the positive effect of risperidone on the NCBRF Conduct Problem subscale was significant in a subgroup of patients with co-occurring “other disorders”, including ADHD. These results should be interpreted with caution as the analyses were unplanned and likely underpowered.
Adolescents (ages 13 to 17)

Evidence on the effectiveness of pharmacotherapy for ADHD in adolescents is very limited (Evidence Tables 1 and 2). We did not find any effectiveness trials or long-term observational studies (assessing functional or safety outcomes) in adolescents with ADHD. Adolescents were studied in one head-to-head trial of MPH IR and SR (OROS) and in 9 placebo-controlled trials of MPH and pemoline. Mixed age populations including adolescents were studied in efficacy trials of atomoxetine, however data are not stratified by school age and adolescents and so are considered in the school-age children section (above).

MPH IR vs MPH OROS (Concerta®)

A single, very small, single blinded crossover study of 6 adolescent boys showed MPH (OROS) superior to MPH IR on some simulated measures of driving skills, dependent on the time of day of testing. ADHD was confirmed using the DePaul ADHD Rating Scale IV (parents completed), the Diagnostic Interview Schedule for Children (DISC-IV), and the Standardized Interview for Adult ADHD. Four of the 6 had inattentive type ADHD. After 7 days of dosing, the teens performed significantly better while taking MPH OROS on 3 of 9 measures (inappropriate braking, missed stop signals, and speed control) at each testing time (2 pm, 5 pm, 8 pm, and 11 pm). Because only F and P-values are reported, it is not possible to interpret the magnitude of differences found. An analysis of a combined score of 7 (of 9) measures at each of the 4 time points indicated that there were no differences between the formulations at the 2 pm and 5 pm test times, but the scores were significantly lower with the IR formulation at the 8 pm and 11 pm times (p< 0.01). Self-evaluations of risky driving behavior did not show any differences between the formulations. Adverse events were not measured. Since 2 teens were previously on MPH OROS, and 2 had been taking MPH IR, and the only person blinded was an observer in the driving simulator, it would be important to know the effect of prior medication and order of randomization. These were not assessed.

MPH IR

Six placebo-controlled crossover trials of MPH IR enrolled a total of 163 adolescents. Patients were diagnosed primarily using the DSM III-R or DSM-IV criteria, however none of the studies reported the ADHD subtypes (with the exception that one reported that all subjects had ‘ADD without hyperactivity’. MPH IR generally was superior to placebo in improving core ADHD symptoms, but was associated with greater frequency of appetite and sleep problems. MPH mean dosages ranged from 8.8 to 75 mg. The trials reported a variety of outcome measures. They were consistent in using various forms of the highly valid Conners’ rating scales (long- and abbreviated forms). However, inconsistency in the way results are reported make estimation of an overall magnitude of effect impossible.

Pemoline

Pemoline reduced ADHD symptoms more than placebo in a small (n = 21) crossover trial of teens with ADHD. Based on reduction of >/= 30% in the ADHD-RS, 60% responded to pemoline, and based on a final CGI of 1 or 2, 55% responded. No significant effects were found with pemoline in another study involving adolescents with comorbid substance use disorder and conduct disorder (see key question 3 for more details). Pemoline was frequently associated with insomnia in both trials.
Indirect Comparison: MPH IR vs Pemoline in Adolescents

Placebo-controlled trials of MPH IR and pemoline in teens do not provide a good opportunity for indirect evidence of comparative efficacy or tolerability due to heterogeneity in outcome reporting. The authors of a pemoline study note that the response rate they found (60%) is lower than seen with methylphenidate (they quote a rate of 73%), but argue that the response rate in teens is less established (than in younger children). The definitions of response compared here are not the same, for example in this study it is a \( \geq 30\% \) change in ADHD-RS score, while the assessment of response in other studies is based on the proportion of variables found significantly different to placebo (e.g. 28 out of 36 dependent variables = 75% response rate). They do provide evidence of a benefit for MPH IR in adolescents, although the degree of benefit or the proportion benefiting is not clear from these studies.

Functional Outcomes: MPH IR

We found extremely limited information on functional capacity outcomes from the clinical trials. Therefore, we included observational studies of \( \geq 6 \) months duration that reported outcomes that reflect functional capacity, for example academic achievement in terms of progression through grades, suicide attempts, police contacts, etc. We found only 2 studies reporting outcomes in adolescents. In an uncontrolled study, a simple follow-up of 16 of 27 (59%) adolescents who had responded to MPH in an uncontrolled study,\(^{115}\) after 6 to 14 months of follow-up the authors simply report that 15 of the 16 had “improved grades”.

In a study using interviews and data from patient charts, 97 young adult males who had taken MPH as children and teens (mean age at discontinuation of MPH was 17 years) were studied.\(^{116}\) There is no comparison group in this descriptive study. The authors conducted a hierarchical analysis to assess the effect of various factors. Significant findings relating to use of MPH were: fewer suicide attempts positively associated with higher dose of MPH and emancipated living situation, and level of relationship commitment were positively associated with response to MPH. Early response to MPH was negatively associated with high school graduation, however.

Adults

Treatment of ADHD in adults has not been widely studied. Only one of three previous, good-quality systematic reviews included studies of adult ADHD.\(^{5}\) There were few studies of only DEX, MPH IR and pemoline in adults available at the time of the Jadad review (1999).\(^{5}\) Jadad et al criticized these studies for their small sample sizes, short durations (\( \leq 6 \) weeks) and for incomplete reporting methods. The review included one study of DEX and MPH\(^{117}\) and placebo-controlled studies of MPH,\(^{118-120}\) pemoline\(^{121}\) and other drugs not included in our review. Jadad et al didn’t draw any conclusions from the study of DEX and MPH because no direct comparisons of these drugs were reported, only changes from baseline.\(^{117}\) They reported that MPH’s efficacy in reducing core ADHD symptoms was inconsistent across placebo-controlled trials and that pemoline was not associated with overall symptom improvement.

Subsequent to the Jadad et al review, other studies have been published that expand the evidence base for DEX,\(^{122-124}\) MPH\(^{125-134}\) and pemoline,\(^{135}\) as well as for amphetamine mixture,\(^{136}\) atomoxetine,\(^{137, 138}\) long-acting forms of bupropion,\(^{125, 135, 139}\) guanfacine\(^{123}\) and modafinil.\(^{122, 140}\) These are included and reviewed here. These studies were fair quality, with one exception.\(^{134}\) The most recent study of MPH is poor quality due to serious concerns about the validity of the outcomes in light of unsuccessful randomization (method not described, but
uneven distribution of age) and uncertainty about characteristics of the groups analyzed (not an ITT).134

**Direct comparisons**

Three head-to-head trials were published subsequent to the Jadad et al review (Evidence Tables 9 and 10).117, 122, 123 All three were designed to evaluate the efficacy of possible alternatives to stimulant therapies.

Results indicate that bupropion SR, guanfacine, and modafinil are viable alternatives to DEX IR and MPH IR (Table 7). Two trials quantified response rates, and found no differences between DEX IR and modafinil (48% vs 48%; p=NS), or between bupropion SR and MPH IR(64% vs 50%; p=NS).125 Differences between trials in treatment duration and method of response rate quantification prevent conclusions about the numerically greater response rates reported for bupropion SR and MPH IR compared to DEX IR and modafinil.125 Bupropion SR and MPH IR response was defined as a CGI score of 1 or 2 (“much” or “very much” improved) and measured over 7 weeks.125 DEX and modafinil were studied for 2 weeks and response was defined as a reduction in ADHD-RS scores of at least 30%.122

Guanfacine and DEX were associated with similar ADHD-RS Total Scores after 2 weeks (22.3 vs 24.2, p=NS) in a crossover study of 17 patients with hyperactive or mixed subtypes of ADHD.123 Improvement magnitude is unknown in absence of mean baseline scores. The endpoint scores remain in the range of clinical significance, however, based on deviance thresholds (+1.5 SD) for ages 30-49 (Total ADHD Score=23.7).141 Neither Bupropion SR or guanfacine offered any advantages over stimulants in reducing depressive symptoms associated with ADHD.123, 125 Nor did guanfacine offer any advantage over DEX in reducing associated anxiety symptoms.123 Bupropion SR,125 guanfacine,123 and modafinil122 were generally similar to stimulants in their effects on cognition.

<table>
<thead>
<tr>
<th>Table 7. Adult ADHD – Comparative trial outcomes</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>N</strong></td>
</tr>
<tr>
<td><strong>Trial</strong></td>
</tr>
<tr>
<td>Kuperman 2001 N=37</td>
</tr>
<tr>
<td>Taylor 2001 N=17</td>
</tr>
<tr>
<td>Taylor 2000 N=22</td>
</tr>
</tbody>
</table>

A: ≥ 30% reduction in ADHD Rating Scale
B: CGI score=1 (“much”) or 2 (“very much”) improved
Indirect comparisons

Placebo-controlled trials were conducted to evaluate whether adults with ADHD benefit from the same treatments that are used in children. Placebo-controlled trials were conducted to evaluate whether adults with ADHD benefit from the same treatments that are used in children.118-121, 124, 126-133, 135-140, 142, 143 Fifty-four percent of trials quantified response rates (Table 8). The rates were variable, even within drug or within method of assessing response: amphetamine mixture=70.4%,136 atomoxetine=52.4%,137 bupropion SR=76%,135 bupropion XL=53%,139 DEX=58%,124 MPH IR=57-78%,119, 120, 128, 133, and pemoline=38.5-50%.121, 143

These placebo-controlled trials fail to provide conclusive evidence about indirect comparative efficacy. It is impossible to attribute differences in response rates to variations in drug effects due to the heterogeneity in design characteristics found across these studies. The wide range of placebo response rates reported (0 – 37%) evidences this heterogeneity.

Sources of heterogeneity include study duration (2-13 weeks), medication dosage levels, response measurement methods, and, most importantly, populations. The primary source of population heterogeneity is variation in ADHD diagnosis methods. Studies differed in ADHD diagnosis methods with regard to usages of diagnostic criteria (Utah criteria, DSM-III-R, or DSM-IV), requirement of second reporter corroboration (i.e., family member), and symptom severity thresholds (e.g., various measurement scale cut-off scores). Studies with more rigorous diagnostic methods122, 128, 133, 139 may be characterized by patients with homogenous symptom presentations; whereas, studies with less stringent criteria120, 121, 124 may be characterized by patients with more heterogenous symptoms.

Other potential sources of population and response differences include distribution of ADHD subtypes and presence of co-existing illnesses. Four studies reported prevalence rates of Inattentive (37-58%), Combined (35-63%) and Hyperactive-Impulsive (0-9%) subtypes.122, 124, 135, 139 Differing subtype prevalence patterns cannot be ruled out in studies that didn’t report this information.119-121, 125, 128, 133, 136, 137, 143 Six studies reported prevalence rates of “any comorbidity” (range=22-57%) and mood/anxiety disorders (range=4.5-68%).119, 120, 128, 135-137 One study focused entirely on patients with ADHD and comorbid cocaine dependence.128 Few studies examined the roles of ADHD subtypes or comorbidities in accounting for drug effects. Those that did reported a lack of adequate statistical power to detect differences and found similar response rates for atomoxetine in patients with inattentive and combined subtypes138 and for atomoxetine and pemoline in patients with comorbidities.137, 143

Table 8. Adult ADHD - Response Rates from Placebo-Controlled Trials

<table>
<thead>
<tr>
<th>Drug</th>
<th>Study Sample Size</th>
<th>Dosage (mean) x Duration (weeks)</th>
<th>Response Rates</th>
</tr>
</thead>
<tbody>
<tr>
<td>Adderall</td>
<td>Spencer 2001</td>
<td>N=27 53.7 mg x 3 70.4% vs 7.4%, p&lt;0.001^</td>
<td></td>
</tr>
<tr>
<td>Atomoxetine</td>
<td>Spencer 1998</td>
<td>N=21 76 mg x 3 52.4% vs 9.5%, p&lt;0.01^</td>
<td></td>
</tr>
<tr>
<td>Bupropion SR</td>
<td>Wilens 2001</td>
<td>N=40 362 mg x 6 76% vs 37%, p=0.02^</td>
<td></td>
</tr>
<tr>
<td>Bupropion XL</td>
<td>Wilens 2005</td>
<td>N=161 393 mg x 8 53% vs 31%, p=0.004^</td>
<td></td>
</tr>
<tr>
<td>DEX</td>
<td>Paterson 1999</td>
<td>N=45 23.85 mg x 6 58% vs 0; p&lt;0.001^</td>
<td></td>
</tr>
<tr>
<td>MPH IR</td>
<td>Wender 1985</td>
<td>N=37 43.2 mg x 2 57% vs 11%; p&lt;0.0001^</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Bouffard 2003</td>
<td>N=30 30-45 mg x 4 63% vs NR, p=NR^</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Schubiner 2002**</td>
<td>N=48 90 mg (max) x 13 77% vs 21%; p&lt;0.05^</td>
<td></td>
</tr>
<tr>
<td>Drug</td>
<td>Study</td>
<td>Dose (mean) x Duration (weeks)</td>
<td>Sample Size</td>
</tr>
<tr>
<td>------------</td>
<td>----------------</td>
<td>-------------------------------</td>
<td>-------------</td>
</tr>
<tr>
<td>MPH IR</td>
<td>Spencer 1995</td>
<td>0.92 mg/kg x 3</td>
<td>N=23</td>
</tr>
<tr>
<td>Pemoline</td>
<td>Wender 1981</td>
<td>71 mg x 6</td>
<td>N=48</td>
</tr>
<tr>
<td></td>
<td>Wilens 1999</td>
<td>148 mg x 4</td>
<td>N=35</td>
</tr>
</tbody>
</table>

A: ≥ 30% reduction in ADHD Rating Scale
B: CGI score=1 (“much”) or 2 (“very much”) improved
C: A or B
D: Moderate to marked improvement on Physician’s Global Rating Scale
E: Other/Unspecified
** Patients comorbid for Cocaine Dependence

Results were contradictory in placebo-controlled trials of MPH that measured core ADHD symptom improvement using a variety of, or unspecified assessment methods.\(^{118,126,127,129}\)

Evidence regarding treatment effects on quality of life and other ADHD-related symptoms (depressed mood, anxiety, and cognition) is not compelling (Table 9). Atomoxetine improved functional status more than placebo (Sheehan Disability Scale Overall Score -4.5 vs -2.9, p=0.022) in one study but this was not replicated in another (-4.4 vs -4.0, p=NS).\(^{138}\) MPH IR consistently reduced associated symptoms of anxiety (HAM-A, Beck Anxiety, POMS)\(^{120,127,133}\) and had a positive effect on cognitive measures (CPT, CPALT),\(^{127,130,131,133}\) regardless of assessment scale used. MPH SR improved depressive symptoms (POMS) after 4 weeks in 37 patients.\(^{129}\)

**Table 9. Adult ADHD – Other symptom-related outcomes in PCTs**

<table>
<thead>
<tr>
<th>Drug</th>
<th>Dose (mean) x Duration (wks)</th>
<th>Sample Size</th>
<th>Effective in treating: Depressive Symptoms</th>
<th>Anxiety Symptoms</th>
<th>Cognition</th>
</tr>
</thead>
<tbody>
<tr>
<td>Atomoxetine</td>
<td>Spencer 1998</td>
<td>76 mg x 3</td>
<td>N=21 48%</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Michelson 2003*</td>
<td>94 mg/day x 10</td>
<td>N=536 65%</td>
<td>No</td>
<td>No</td>
<td>No</td>
</tr>
<tr>
<td>Bupropion SR</td>
<td>Wilens 2001</td>
<td>362 mg/day x 6</td>
<td>N=40 55%</td>
<td>No</td>
<td>No</td>
</tr>
<tr>
<td>Bupropion XL</td>
<td>Wilens 2005</td>
<td>393 mg x 8</td>
<td>N=161 60%</td>
<td>No</td>
<td>No</td>
</tr>
<tr>
<td>DEX</td>
<td>Paterson 1999</td>
<td>24 mg/day x 6</td>
<td>N=45 60%</td>
<td>No</td>
<td>No</td>
</tr>
<tr>
<td>MPH IR</td>
<td>Bouffard 2003</td>
<td>30-45 mg (tid) x 4</td>
<td>N=30 80%</td>
<td>No</td>
<td>Yes (HAM-A)</td>
</tr>
<tr>
<td>Gualtiero 1985</td>
<td>0.3 mg/kg (bid) x 5 days</td>
<td>N=8 100%</td>
<td>-</td>
<td>-</td>
<td>Yes (CPT)</td>
</tr>
<tr>
<td>Kinsbourne 2001</td>
<td>5, 10, or 20 mg/day (QD) x 1 day</td>
<td>N=17 41%</td>
<td>-</td>
<td>-</td>
<td>Yes (CPALT)</td>
</tr>
<tr>
<td>Tenenbaum 2002</td>
<td>45 mg/d (max; qid) x 3</td>
<td>N=24 46%</td>
<td>No</td>
<td>Yes (Beck Anxiety)</td>
<td>Yes (CPT)</td>
</tr>
<tr>
<td>MHender 1985</td>
<td>43 mg x 2</td>
<td>N=37 54%</td>
<td>Yes (POMS)</td>
<td>Yes (POMS)</td>
<td>-</td>
</tr>
<tr>
<td>MPH SR</td>
<td>Levin 2002</td>
<td>20 mg/day x 4</td>
<td>N=347 63%</td>
<td>Yes (POMS)</td>
<td>-</td>
</tr>
<tr>
<td>Modafinil</td>
<td>Turner 2004</td>
<td>200 mg x single dose</td>
<td>N=20 65%</td>
<td>-</td>
<td>-</td>
</tr>
</tbody>
</table>

*2 studies reported together
Functional Outcomes: MPH IR

We found extremely limited information on functional capacity outcomes from the clinical trials. Therefore, we included observational studies of \( \geq 6 \) months duration that reported outcomes that reflect functional capacity, for example academic achievement in terms of progression through grades, suicide attempts, police contacts, etc. We found 2 studies that reported these outcomes among adult patients who had been treated as children \(^{70, 115, 116, 144, 145}\). Due to various methodological limitations, these studies do not provide good evidence for long-term effectiveness even for MPH.

In a cross-sectional follow-up study of young men diagnosed with ‘persistent hyperactivity’ at ages 6 to 12 years, those who had not received medication were compared to a group that had received MPH for at least 3 years during childhood.\(^ {145}\) The groups were initially seen in different time-periods, separated by 5 to 15 years. Because the groups were from different periods, a third group of normal children who were contemporaneous to the MPH group was added. The sizes of the groups also differed, with 64 in the non-treated hyperactive group, 20 in the MPH treated group, and 20 in the normal controls and data were not available for all subjects on all questions. Mean follow-up of the hyperactive groups was 10 to 12 years. No information on baseline characteristics from childhood is given. No consistent differences in functional outcomes were found between the MPH and untreated groups (Table 10). Considering the potential confounding of differences in the years the children were treated, and the very small numbers of subjects per group per variable, these results should be interpreted with caution.

### Table 10. Long-term functional outcomes of MPH from Hechtman 1984

<table>
<thead>
<tr>
<th>Variable</th>
<th>Favors</th>
<th>MPH group</th>
<th>Non-treated</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age at follow-up</td>
<td>NA</td>
<td>22 years</td>
<td>20 years</td>
<td>&lt;0.01</td>
</tr>
<tr>
<td>Living with girlfriend/wife (n)</td>
<td>MPH</td>
<td>8</td>
<td>5</td>
<td>&lt;0.01</td>
</tr>
<tr>
<td>Duration last job held</td>
<td>Non-treated</td>
<td>21 weeks</td>
<td>70 weeks</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Aggression</td>
<td>Untreated</td>
<td></td>
<td></td>
<td>&lt;0.06</td>
</tr>
<tr>
<td>Psychiatric treatment at present</td>
<td>MPH</td>
<td>1</td>
<td>22</td>
<td>&lt;0.02</td>
</tr>
<tr>
<td>Age starting alcohol use</td>
<td>Non-treated</td>
<td>14.8 years</td>
<td>16.2 years</td>
<td>&lt;0.03</td>
</tr>
<tr>
<td>Duration of alcohol use</td>
<td>Non-treated</td>
<td>25 months</td>
<td>10.8 months</td>
<td>&lt;0.05</td>
</tr>
<tr>
<td>Abuse/addiction to alcohol (n)</td>
<td>MPH</td>
<td>13</td>
<td>26</td>
<td>&lt;0.05</td>
</tr>
<tr>
<td>Age at first cocaine use</td>
<td>MPH</td>
<td>20 years</td>
<td>18.9 years</td>
<td>&lt;0.02</td>
</tr>
<tr>
<td>Age stopping cocaine use</td>
<td>Non-treated</td>
<td>22 years</td>
<td>18.9 years</td>
<td>&lt;0.001</td>
</tr>
</tbody>
</table>

The MPH group in this study was previously reported after 5 years of follow-up (as adolescents), with comparison groups of boys treated with chlorpromazine or untreated boys.\(^ {144}\) This study reported academic performance, with no differences found between the groups.

Drugs with No Evidence for the Treatment of ADHD

*No trial evidence for the treatment of ADHD was found for Aripiprazole, Clozapine, Olanzapine, Quetiapine, or Ziprasidone in any age group.*
Key Question 2: What is the comparative tolerability and safety of different pharmacologic treatments for attention deficit disorders?

Short-Term Trial Evidence in Young Children (Preschool Age; 3-5 years)

One of three placebo-controlled trials of MPH IR reported results of adverse event assessments.\(^{50}\) MPH IR was clearly associated with higher rates of increased sadness, decreased appetite, and sociability impairments than placebo after 7-10 days in 31 preschoolers.

Short-Term Trial Evidence in Children (Elementary School Age: 6-12 years)

Adverse events were reported in 18 (of 39) head-to-head trials. The results are summarized in Table 11 below, full reporting of adverse event data can be found in Evidence Table 3.

Stimulants

Four of 6 trials of DEX versus MPH IR reported no differences between the drugs in adverse events.\(^{34, 64-66}\) However, 2 short-term crossover trials found DEX to cause greater weight loss than MPH IR with mean weight change differences of 0.7 kg to 0.97 kg.\(^{44, 67}\) One of 3 trials of Amphetamine mixture versus MPH IR found no difference in adverse event rates,\(^{82}\) but 2 other studies found differences.\(^{42, 81}\) Limitations in study design and lack of description of analysis methods make results from these studies less reliable. These studies found that adding additional doses to the daily regimen of either drug increased the reports of loss of appetite and sleep problems,\(^{81}\) and that Amphetamine mixture given twice daily caused the highest rates of these adverse events.\(^{81}\) All 3 studies of MPH IR versus extended release formulations (MPH OROS, MPH SODAS, and MPH SR) that reported no significant differences in the incidence of side effects.\(^{57, 59, 60}\) Amphetamine mixture and DEX SR were found to cause more weight loss than DEX IR during the first week of treatment, but weight gain during the second week was greater with these drugs than with DEX IR.\(^{84}\) Since this was such a short-term trial, no conclusions about differential effects on weight can be made from these data. No differences in adverse event rates were found between MPH SR (Ritalin LA®) and MPH OROS (Concerta®).\(^{38}\)

Non-stimulants versus MPH IR

Clonidine was found to have significantly higher rates and greater severity of sedation among children with both ADHD and Tourette’s Disorder.\(^{46}\) There were no differences in the severity of tics between the groups. No differences in adverse events were found between bupropion IR and MPH IR.\(^{24}\) Atomoxetine caused significantly more vomiting and somnolence in one study\(^{89}\) and paresthesias in the other\(^{27}\) while MPH IR caused more ‘abnormal thinking’.
### Table 11. Summary of Adverse Effects Reported

<table>
<thead>
<tr>
<th>Study</th>
<th>Differences Found</th>
<th>Study</th>
<th>No Differences Found</th>
</tr>
</thead>
<tbody>
<tr>
<td>DEX versus MPH IR</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Kauffman 1981</td>
<td>Significant difference found only on weight change:</td>
<td>Arnold 1978</td>
<td>P = NS (incidence and weight change)</td>
</tr>
<tr>
<td>Crossover; N=29</td>
<td>Mean change in weight (kg): DEX -0.86 vs MPH +0.11 (Difference 0.97kg, p NR)</td>
<td>Efron 1997</td>
<td>P = NS for all (incidence only)</td>
</tr>
<tr>
<td>6 weeks</td>
<td></td>
<td>Crossover; N=125</td>
<td></td>
</tr>
<tr>
<td>Sharp 1999</td>
<td>Mean change in body weight (kg) reported to be greater with DEX. Difference 0.7 kg</td>
<td>Elia 1991</td>
<td>P = NS (incidence and severity)</td>
</tr>
<tr>
<td>Crossover; N=32</td>
<td>Dextroamphetamine: -1.1; p=0.01 from baseline</td>
<td>Elia 1993</td>
<td>P = NS (incidence)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Crossover; N=33</td>
<td></td>
</tr>
<tr>
<td>Adderall versus MPH IR</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pelham 1999a</td>
<td>Authors assessment notes that adding an afternoon dose of either drug resulted in increased reports of loss of appetite or sleep delay. Statistical comparison NR.</td>
<td>Pliszka 2000 Paralle; N=58</td>
<td>All p=NS (incidence only)</td>
</tr>
<tr>
<td>Crossover; N=21</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pelham 1999b</td>
<td>Authors note that differential side effects were seen only for loss of appetite and trouble sleeping with the high (12.5mg/day) dose of Adderall. (p-values NR).</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Crossover; N=25</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>IR versus SR formulations of MPH</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pelham 2001</td>
<td></td>
<td>P = NS ((incidence only)</td>
<td></td>
</tr>
<tr>
<td>Crossover; N=70</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Wolraich 2001</td>
<td></td>
<td>P = NS (incidence only)</td>
<td></td>
</tr>
<tr>
<td>Parallel; N=312</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Whitehouse 1980</td>
<td></td>
<td>P = NS (incidence only)</td>
<td></td>
</tr>
<tr>
<td>Parallel; N=34</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Clonidine vs MPH IR</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Tourette's Study Group 2002</td>
<td>Clonidine vs MPH IR</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Parallel; N=136 8%; p=0.007</td>
<td>Sedation (% patients): 48% vs 14%; p=0.004</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Sedation (% patients rated as moderate or severe): 35% vs 20%</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Other comparisons to MPH IR</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Kragtchvili 2002</td>
<td>Atomoxetine vs MPH IR; p=NS on 24 of 27 AEs reported;</td>
<td>Barrickman 1995 Crossover; N=18</td>
<td>P = NS (incidence only)</td>
</tr>
<tr>
<td>Parallel; N=228</td>
<td>Atomoxetine worse:</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Vomiting: 22 (12%) vs 0, p=0.017</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Buitelaar1996</td>
<td>Atamoxetinone worse:</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Crossover; N=32</td>
<td>Somnolence: 20 (10.9%) vs 0, p=0.029</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Thinking abnormal: 0 vs 2 (5%); p=0.031</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Multiple Comparisons</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>James 2001</td>
<td>Based on SERS assessment tool: ANOVA analysis indicates Adderall and DEX SR caused greater decreases in weight than DEX IR, however these groups also had greater recovery of weight during the 2nd week (compared to DEX IR in each case). All other findings p = NS for drug vs drug comparisons.</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Growth Effects

A study of withdrawing MPH IR during summer months versus not withdrawing assessed the effect on weight and height. Children with cross-situational, pervasive hyperactive behavior (n = 62) were randomized and followed for a 3 year period. Overall, 42% of those randomized withdrew, with data available for 58 children at the end of summer 1 (ON n=32, OFF n=26); and 34 at the end of summer 2 (ON n=20, OFF n=14). Weight and height were collected by unblinded secretaries, but not for the purposes of this study. Both groups gained in weight and height over each summer, but during summer 1, the MPH IR ON group gained significantly less (0.9 kg, p=0.005) than the MPH IR OFF group. However, in summer 2 the difference was non-significant (0.6 kg). The effect on height was the reverse of these finding, with no significant difference in summer 1 (0.1 cm), but a significant difference after summer 2 (1.3 cm, p=0.02). The serious limitations of this study, in design and conduct, limit the likelihood that the findings are valid.

Adolescents

Placebo-controlled trials of MPH IR and pemoline provide limited evidence of short-term stimulant tolerability in adolescents. MPH IR was associated with significant appetite and sleep disturbances across some, but not all placebo-controlled trials. Additionally, adolescents taking MPH IR frequently reported increases in dulled affect, social withdrawal, irritability and stomachache in two placebo-controlled trials. Adolescents taking pemoline in two trials also reported insomnia at a greater frequency than those taking placebo. Other significant adverse events associated with pemoline included loss of appetite, stomachaches and “picking” at skin. Indirect comparisons of MPH IR and pemoline are not possible due to heterogeneity in adverse event reporting.

Adults

There is considerable interest in alternative, nonstimulant treatments for ADHD to address the needs of individuals intolerant of adverse effects that are often associated with stimulants (e.g., insomnia, appetite suppression). Therefore, this review particularly addresses the important question of whether current nonstimulant alternatives do, indeed, offer such an advantage over stimulants.

In summary, randomized controlled trials do not provide evidence that any one stimulant is more tolerable than another or that nonstimulants are more tolerable than stimulants. Trials were short-term in duration and heterogenous for types of adverse events measured. Adverse events were inadequately defined and ascertainment methods were unclear.

Direct comparisons of stimulants vs nonstimulants

Bupropion SR, modafinil or guanfacine offered no clear advantage over MPH IR or DEX IR in terms of preventing common adverse effects often associated with stimulants. There were no differences between stimulants and nonstimulants in muscle tension (ranging 24-29.4% for stimulants and 19% for bupropion SR), insomnia (ranging 16.7-38% for stimulants and 15.4-19% for non-stimulants), appetite suppression (ranging 24-25.4% for stimulants and 19% for modafinil), and headache (10% DEX IR and 10-15.4% for non-stimulants).

Bupropion SR and MPH were associated with similar rates of overall adverse events (69% vs 75%) and withdrawals due to adverse events (0 vs 16.7%) were not statistically different. However, the reason for not finding a difference may have been a lack of statistical
power due to inadequate numbers of subjects this trial of 37 participants.\textsuperscript{125} Overall tolerability profile data for modafinil or guanfacine compared to DEX was not reported.\textsuperscript{122, 123}

**Indirect comparisons**

Heterogeneity in design characteristics (discussed above) and in adverse effects measurement and reporting methods prevents reliable indirect comparisons of tolerability across placebo-controlled trials of stimulants and nonstimulants (Evidence Table 11). Table 12 reflects placebo-controlled trials that reported rates of adverse events commonly associated with stimulant therapy, as well as withdrawals. The rates of insomnia and appetite loss in the placebo groups in these trials varied widely. Within trial differences between drug and placebo group were only found in 3 instances: atomoxetine was associated with insomnia and appetite loss significantly more often based on a pooled analysis of two trials, amphetamine mixture was found to have significantly higher rate of appetite loss, and MPH IR was found to have a significantly higher rate of insomnia (Table 12).\textsuperscript{136, 138} Overall, more patients taking atomoxetine withdrew due to adverse events than in the placebo group (8.5\% vs 3.4\%; \(p\)=0.03). These findings suggest that atomoxetine may not offer an advantage in tolerability over stimulants with regard to insomnia and appetite loss. Bupropion XL was associated with nonsignificant increases in insomnia and withdrawals due to adverse events. However, it may be that the lack of finding a significant difference with the other drugs was due to inadequate sample sizes.

<table>
<thead>
<tr>
<th>Trial</th>
<th>N</th>
<th>Mean age (yrs) % male</th>
<th>Treatment (mean) x Duration (wks)</th>
<th>Insomnia</th>
<th>Appetite Loss</th>
<th>Withdrawal due to AEs</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Stimulants</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Spencer 2001\textsuperscript{128}</td>
<td>N=27</td>
<td>38.8 years 56%</td>
<td>Amphetamine mixture 53.7 mg x 3</td>
<td>37 vs 14.8% (ns)</td>
<td>29.6 vs 11.1% (p=0.03)</td>
<td>-</td>
</tr>
<tr>
<td>Schubiner 2002\textsuperscript{128}</td>
<td>N=48</td>
<td>37.5 89.6%</td>
<td>MPH IR 90 mg x 3</td>
<td>63% vs 33%; (p&lt;0.05)</td>
<td>50% vs 25% (ns)</td>
<td>0 vs 4.2% (ns)</td>
</tr>
<tr>
<td>Wender 1981\textsuperscript{121}</td>
<td>N=48</td>
<td>28.3 years 54%</td>
<td>Pemoline 71 mg (qd in am) x 6 weeks</td>
<td>6 vs 0 cases</td>
<td>-</td>
<td>15.4% vs 0 (ns)</td>
</tr>
<tr>
<td>Wilens 1999\textsuperscript{43}</td>
<td>N=35</td>
<td>40.7 years 68.6%</td>
<td>Pemoline 148 mg (qd, schedule nr) x 4 weeks</td>
<td>-</td>
<td>8% vs 0</td>
<td>11.4% vs 0</td>
</tr>
<tr>
<td><strong>Nonstimulants</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Michelson 2003\textsuperscript{18a}</td>
<td>N=536</td>
<td>41.1 years 65%</td>
<td>Atomoxetine bid 94.4 mg x 10</td>
<td>20.8 vs 8.7% ((p&lt;0.001))</td>
<td>11.5 vs 3.4% ((p&lt;0.001))</td>
<td>8.5% vs 3.4%; (p=0.03)</td>
</tr>
<tr>
<td>Wilens 2005</td>
<td>N=162</td>
<td>40.2 59.9%</td>
<td>Bupropion XL 393 mg x 8</td>
<td>12% vs 7% (ns)</td>
<td>-</td>
<td>4.9% vs 0 (ns)</td>
</tr>
</tbody>
</table>

* pooled results from 2 trials
Evidence on the Long-Term Safety of drugs used to treat ADHD

We included 18 observational studies for analysis of long-term safety parameters.\textsuperscript{131, 148-164} Seven studies used cohort designs to compare groups taking MPH to DEX,\textsuperscript{153, 154, 161, 163} imipramine,\textsuperscript{151} unmedicated hyperactives,\textsuperscript{161, 162} and general population growth chart norms.\textsuperscript{152} Eleven non-comparative studies involved patients exposed to MPH IR,\textsuperscript{131, 149, 150, 156, 157, 159} MPH SR (OROS),\textsuperscript{158} atomoxetine,\textsuperscript{148, 153, 164} or Adderall.\textsuperscript{160} All but two studies were 1-5 years in duration.\textsuperscript{131, 149} All but one study involved elementary school-aged children.\textsuperscript{160} The exception was one before-after study of amphetamine mixture in adults with ADHD.\textsuperscript{160}

Growth (height and weight) was commonly reported in these studies. Other long-term safety outcomes were assessed, including tics, seizures, cardiovascular adverse events, injuries, and attempted suicide.

No study was rated good quality. All but one was rated fair quality due to biased patient selection processes and/or biased or unspecified outcome ascertainment methods. We did not analyze results from a poor-quality, comparative study of growth rebound in MPH and DEX due to our concerns about how possible additional biases may have affected the results.\textsuperscript{163} We cannot rule out the possibility of between groups differences in baseline characteristics because no information/analysis was provided. We also cannot rule out the possibility that the results were confounded by time and other relevant factors.

Height and weight effects

A frequently cited nonsystematic review concluded that effects on weight and height associated with MPH IR vary across short-term clinical trials and long-term observational studies and are mostly transient.\textsuperscript{165} We reached similar conclusions based on our analysis of a larger number of primarily long-term observational studies that compared MPH IR to DEX IR,\textsuperscript{154, 155, 161} imipramine,\textsuperscript{151} or unmedicated hyperactive control groups.\textsuperscript{157, 161, 162} Height and weight changes associated with MPH IR\textsuperscript{150, 152, 156, 158, 159} and OROS\textsuperscript{158} or atomoxetine\textsuperscript{164} were also observed in long-term noncomparative studies.

Comparative studies

\textit{Height.} These studies do not answer the question of whether any one stimulant suppresses growth in height any more than any other, nor do they clearly support a relationship between MPH and suppression of height.

The only comparative evidence comes from two studies of DEX and MPH.\textsuperscript{154, 161} Results are mixed across these studies (Table 13). Both reported changes in height percentiles using the outdated Iowa City norms. DEX and MPH were both associated with similar height \textit{increases} at final follow-up (mean 6 years) in one study,\textsuperscript{154} and DEX was associated with significantly greater height \textit{decreases} than MPH after at least two years in the other.\textsuperscript{161} It is impossible to establish whether heterogeneity in group characteristics across studies may possibly contribute to the contradictory findings, as one of the studies did not report mean age, dosage, or duration.\textsuperscript{161}

\textit{Weight.} Results from three comparative studies suggest that DEX is associated with significantly greater suppression of weight gain than MPH, at least in the first 1-2 years (Table 13).\textsuperscript{154, 155, 161} DEX was associated with a significantly lower mean weight gain (kg) than MPH after nine months in one study\textsuperscript{155} and significantly greater declines in weight percentiles after the first of 5 years in one study\textsuperscript{154} and at end of treatment (≥ 2 years) in another.\textsuperscript{161} In the 5-year, partly retrospective and partly prospective study that involved 84 children (mean age at initiation...
of drug therapy=9 years and 82% male), however, differences in decreased weight percentiles between DEX and MPH resolved by the second year and resulted in significantly greater than expected mean increases in weight percentiles at final follow-up (+10.9, p<0.01 and +12.8, p<0.001, respectively).154

The 9-month study also reported a few subgroup analyses.155 The first suggests that comparison of mean weight gain between DEX and MPH may have been confounded by dosage disparities. Apparently, the difference between DEX and MPH resolved when four patients taking lower-dose MPH (20 mg/day) were removed from the analysis (0.13 vs 0.12 kg per month). Also, weight gain in children who continued medication over the summer versus those who discontinued medication during the summer was also reported. In patients taking DEX, medication continuation was associated with significantly lower mean weight gain than in children who discontinued medication (0.14 vs 0.47 kg per month, p<0.01). Medication continuation status did not have an effect on weight gain in the group of patients taking MPH. MPH was associated with decreases in weight percentiles similar to imipramine after one year151 and absolute weight changes that were similar those in unmedicated healthy controls in another 2-year study.157 Results were mixed across two studies that compared children taking MPH to unmedicated hyperactives, however.151,162 In one study, MPH was associated with significantly greater declines in weight percentiles than in the unmedicated children after one year.151 The differences between the MPH groups and the unmedicated group increased numerically along with the dosages (< 20 mg= -6.88, 20.56 mg= -8.81, > 20 mg=-15.40, all p<0.005). In the other study, the MPH group and the unmedicated group demonstrated similar absolute weight gain (kg) after 364 days.162

Table 13. Long-Term Height and Weight Outcomes in Observational Studies

<table>
<thead>
<tr>
<th>Study</th>
<th>Interventions (mean dose) x duration</th>
<th>Age Population</th>
<th>Height</th>
<th>Weight</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gross 1976</td>
<td>DEX 16.5 mg, n=12 (average follow-up=6.8 yrs)</td>
<td>Mean age=9 82% male Children/adolescents with hyperkinetic syndrome or minimal brain dysfunction</td>
<td>Change in percentile: +10.9, p&lt;0.01 vs +12.8, p&lt;0.001</td>
<td>Change in percentile: +16.0, p&lt;0.02 vs +11.4, p&lt;0.001</td>
</tr>
<tr>
<td></td>
<td>MPH 34 mg, n=60 (average follow-up=5.8 years)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Safer 1972</td>
<td>MPH 37.5 mg, n=4/24.0 mg, n=5 DEX 11.7 mg, n=3/11.8 mg, n=8 x 9 months</td>
<td>Mean age=9.8 Gender NR</td>
<td></td>
<td>Weight gain (kg): 0.23 vs 0.12, t=1.8, p&lt;0.05 Weight gain (excluding patients taking low-dose MPH, n=16) (kg): 0.13 vs 0.12, t=0.137, NS ON vs OFF Weight gain (kg) over a 3-month summer period: MPH= 0.29 vs 0.41, t=0.526, p=NS; DEX= 0.14 vs 0.47, t=2.523, p&lt;0.01</td>
</tr>
</tbody>
</table>
### Studies

<table>
<thead>
<tr>
<th>Study</th>
<th>Interventions (mean dose) x duration</th>
<th>Age</th>
<th>Height</th>
<th>Weight</th>
</tr>
</thead>
<tbody>
<tr>
<td>Safer 1973</td>
<td>DEX, n=29 MPH, n=20 Unmedicated controls, n=14 x ≥ 2 years Mean dosages NR</td>
<td>Mean age nr 89.8% male in children on medication; 100% male in unmedicated control group 100% white</td>
<td>Change in percentile points: DEX: -13.45 MPH high-dose (&gt; 20 mg): -9.40 All MPH: -5.20 MPH low-dose (≤ 20 mg): -1.00 Controls: +1.29</td>
<td>DEX; MPH: high-dose (&gt; 20 mg), all, low-dose (≤ 20 mg); controls Percentile changes in: Weight: -20.38; -10.0; -6.35; -2.7; +6.79 DEX &gt; all MPH dosage groups and controls; MPH high-dose and all doses vs controls; MPH low-dose&gt;controls</td>
</tr>
<tr>
<td>Quinn 1975</td>
<td>Condition 1: Imipramine 65.4mg, n=12 Condition 2: MPH 20.56 mg, n=23 Condition 3: MPH &gt; 20 mg a day (n=5) Condition 4=MPH 20 mg a day or less (n=18) Condition 5=No treatment, n=12 x 1 year</td>
<td>Mean age nr 100% male Hyperactivity</td>
<td>Height changes (percentile scores): -2.20 vs +3.19 vs -3.0 vs +5.12 vs -1.46 t-scores for comparisons of condition 5 with 1; 2; 3; 4 (p-values all NS): 0.23; 1.05; 0.22; 1.59 t-scores, p-values for comparisons of condition 1 with 2, 3, and 4: 1.25, p=NS; 0.12, p=NS; 1.90, p&lt;0.05</td>
<td>Weight change (percentile): -7.54 vs -8.81 vs -15.40 vs -6.88 vs +1.61 t-scores, p-values for comparisons of condition 5 with 1; 2; 3; 4: 2.45, p&lt;0.01; 3.42, p&lt;0.005; 4.18, p&lt;0.005; 3.44, p&lt;0.005 t-scores, p-values for comparisons of condition 1 with 2; 3; 4: .37, p=NS; 1.27, p=NS; 0.19, p=NS</td>
</tr>
<tr>
<td>McNutt 1976</td>
<td>MPH vs nonmedicated vs controls 12 months: n=28, n=24, n=47 24 months: n=13, n=10, n=14</td>
<td>12-month 10.2 years 77.8% male 24-month 9.9 years 86.5% male</td>
<td>Height change (cm): +12.1 vs +12.1 Height at end of treatment (cm): 150.4 vs 148.3; p=NS</td>
<td>Weight change (kg): 42.0 vs 40.3; p=NS</td>
</tr>
<tr>
<td>Zeiner 1995</td>
<td>Medicated (MPH 23 mg) vs unmedicated x 1.7 years</td>
<td>Hyperactivity mean age 9.0 yrs 100% male Ethnicity NR</td>
<td>End of study measurements: Weight (kg): 42.0 vs 40.3; p=NS</td>
<td></td>
</tr>
</tbody>
</table>

### Noncomparative Studies

Multiple noncomparative study findings provide inconclusive evidence regarding MPH IR effects on children’s height and weight. A pooled analysis of data from open-label extensions of 13 trials of atomoxetine assessed the effect on height and weight.148

**Height.** In summary, studies of children taking MPH IR at various doses for 1-4 years showed inconsistent suppression of growth in height as compared to children taking imipramine,151 those who were unmedicated,151, 157, 162 and in noncomparative studies that reported varied analyses including differences between expected and actual growth,150 change in percentile,152 percent of expected growth,156 and proportion of patients with decreased growth.
rates. One before-after study followed 407 children with ADHD (mean age=9.2 years and 83% male) taking MPH SR (OROS) 41 mg/day for 12 months. MPH SR was associated with a steady increase in mean absolute height of 5.2 cm. No comparisons to norms or controls were made.

**Weight.** Noncomparative studies provide mixed evidence about the association between MPH and suppression of weight gain. In the earliest study (1977), only 2 of 36 boys with minimal brain dysfunction (5.5%) gained weight while taking MPH (max dose 20 mg) over 16 months. The other 34 boys gained weight. The next study, published in 1979, involved 72 boys (age range 6-12) with hyperactivity that were taking MPH for up to two years. A significant growth weight deficit (30%, p<0.05) was associated with MPH 24.2 mg/day (0.47 mg/kg) in the 72 boys who completed the first year. The growth weight deficit associated with MPH 0.59 mg/kg of 10% was insignificant for the 48 boys who completed the second year of treatment. Results of a subgroup analysis suggests that the deficit in weight gain was only significant in patients that continue to use medication over the summer months compared to those who did not. The third study, published in 1983, involved relatively higher mean dosages of MPH (39.9 to 41.3 mg) and followed children with hyperactivity over the longest observation period (4 years). MPH was associated with significant declines in weight percentiles in all four years of the study (Years 1: -9.7 vs 2: -15.9 vs 3: -18.6 vs 4: -20.8; p<0.001 for all). The final study, published in 1999, found an insignificant difference (0.72 kg) between expected vs actual weight gain in 29 patients who took MPH 34.5 mg for two years.

One before-after study followed 407 children with ADHD (mean age=9.2 years and 83% male) taking MPH SR (OROS) 41 mg/day for 12 months. MPH OROS was associated with a steady increase in mean absolute weight of 2.6 kg.

Based on 412 patients (children and adolescents) who had received atomoxetine for at least 2 years and had at least one post-baseline height and weight measurement, atomoxetine resulted in a mean decrease in expected weight of 0.87 kg, and decrease in expected height of 0.44 cm. Analysis of change over time indicated that weight changes were greatest in the early months of treatment, with some regression toward the mean percentile at 2 years. Height changes appeared to occur over a longer period of time, but also regressed toward the mean by 2 years. Results from another before-after study of 10 boys (mean age nr) suggested that atomoxetine was associated with a weight loss of 1.15 kg after 10 weeks.

**Tics**

Four studies reported tic-related outcomes. One of these is a long-term placebo-controlled trial of MPH IR. Although the 1-year study started out with similar numbers assigned to placebo and MPH, but the study end 72 were on MPH and only 18 on placebo. Development of new tics or worsening of pre-existing tics was not different between the two groups. Two of the observational studies involved children and the other involved adults. These studies do not provide any information about how different pharmacologic treatments for ADHD compare in safety with regard to tic-related outcomes. Table 14 summarizes the characteristics and outcomes from these studies.
Table 14. Tic-related outcomes in observational studies

<table>
<thead>
<tr>
<th>Study</th>
<th>Interventions (mean dose)</th>
<th>Sample size</th>
<th>Population</th>
<th>Age</th>
<th>Gender</th>
<th>Tics</th>
</tr>
</thead>
<tbody>
<tr>
<td>Law 1999</td>
<td>MPH 0.5 mg/kg twice daily vs Placebo N = 72</td>
<td>ADHD with no prior treatment for tics or ADHD</td>
<td>8.4 years</td>
<td>76% males</td>
<td>1 year</td>
<td>New onset tics: 19.6% MPH IR vs 16.7% placebo (NS) Exacerbation of pre-existing tics: 33% both groups (NS)</td>
</tr>
<tr>
<td>Gadow 1999</td>
<td>MPH 34.5 mg N=29</td>
<td>ADHD and chronic tics</td>
<td>8.8 years</td>
<td>91.2% male</td>
<td>2 years</td>
<td>Insignificant changes in frequency/severity of motor/vocal tics</td>
</tr>
<tr>
<td>Wilens 2003</td>
<td>MPH OROS 41 mg/day N=407</td>
<td>ADHD</td>
<td>9.2 years</td>
<td>83% male</td>
<td>1 year</td>
<td>New onset tics (% patients): 23 (6.4%)</td>
</tr>
<tr>
<td>Horrigan 2000</td>
<td>Adderall – 10 mg/day, titrated (mean NR) N=24</td>
<td>ADHD</td>
<td>33 years</td>
<td>50% male</td>
<td>1 year</td>
<td>Motor tics (% patients): 1 (4%)</td>
</tr>
</tbody>
</table>

Seizures

The study that compared MPH (< 20 mg/day, 20.56 mg/day, and > 20 mg/day) to imipramine 65.4 mg/day and an untreated group (discussed earlier) also assessed seizures as an adverse event. None of the 70 males with hyperactivity experience a seizure over the one-year study period.

Cardiovascular adverse events

One study involved 169 children and adolescents that continued on open or blinded atomoxetine (max dose of 2 mg/kg divided into twice daily) for at least 1 year following 3 short-term, placebo-controlled trials. The timing of ECG measurements is not stated, but is presented by increasing dose. Linear regression suggests that there is no evidence of an increase in QTc with increasing dosage of atomoxetine.

Substance Abuse

Two studies assessing the risk of substance abuse among children treated with MPH or not treated, with similar duration of follow-up (mean just under 2 years) found no differences between the groups in the risk of using cigarettes, marijuana, or alcohol. A questionnaire – based survey reported small numbers of teenagers who reported having taken higher than prescribed doses, purposefully mixing ADHD drugs with other substances, having lost a bottle of ADHD medications, etc. Multiple ADHD medications had been used by the survey respondents, and these results do not provide insight into comparative risk for future substance abuse among users of ADHD drugs.

Injuries

A retrospective database study analyzed an association between childhood behavioral disorders and common childhood injuries by using the British Columbia Linked Health Data Set to identify injuries. Children with behavioral disorders were identified using MPH prescriptions as a proxy for diagnosis using data in a Triplicate Prescription Program. Injury frequencies in children prescribed MPH at least once between 1/1/1990 and 12/31/1996 (n=16,806) were compared to those in children not taking MPH (n=1,010,067). Children were 51.4% male and
aged less than 19 years. Mean duration of exposure was not identified. Odds of any injury (fractures, open wounds, poisoning/toxic effect, intracranial, concussion, and burns) were significantly higher in children taking MPH than for those not taking MPH (OR 1.67, 95% CI 1.54 to 1.81), even after adjusting for baseline age, sex, socioeconomic status and region. This study design clearly suffers from lack of sensitivity to diagnosis, in that an unknown number of children with behavioral disorders are included in the group not taking MPH. Since MPH was used simply as a proxy for behavioral disorders, the relationship between the drug and the increase in injuries is not necessarily clear.

**Suicide**

One before-after study followed 8 adult males (mean age of 27.2 years) that continued on open MPH for three to six months subsequent to participation in short-term clinical trials. One participant (12.5%) attempted to commit suicide by consuming a month’s supply of MPH.

**Removal from Market or FDA Warnings**

The information provided below does not come from comparative studies but is included here to provide a complete and current picture of important safety concerns.

**Amphetamine Mixture SR (Adderall XR®)**. Health Canada suspended the market authorization of ADDERALL XR® in February 2005. Health Canada's decision was a result of 20 international reports of sudden death in patients taking either Adderall or Adderall XR (Adderall IR was not marketed in Canada). FDA review of the reports of sudden deaths in children resulted in the following statement: “SUD (sudden unexplained death) has been associated with amphetamine abuse and reported in children with underlying cardiac abnormalities taking recommended doses of amphetamines, including Adderall and Adderall XR. In addition, a very small number of cases of SUD have been reported in children without structural cardiac abnormalities taking Adderall. At this time, FDA cannot conclude that recommended doses of Adderall can cause SUD, but is continuing to carefully evaluate these data.” ([http://www.fda.gov/cder/drug/infopage/adderall/default.htm](http://www.fda.gov/cder/drug/infopage/adderall/default.htm)). In August of 2005, a specially convened committee, highlighting the potential risk for stimulants in general, recommended that Health Canada reinstate the licensure of Adderall XR®, with modifications to the drug label, including: warnings that sudden death has occurred at therapeutic doses in children with structural cardiac abnormalities and that Adderall XR should generally not be used in patients with pre-existing structural cardiac abnormalities. Current FDA labeling includes a boxed warning about abuse potential and that “misuse of amphetamine may cause sudden death and serious cardiovascular adverse events.”

**Atomoxetine**. Two case reports (via the FDA MedWatch system) of hepatotoxicity in patients taking atomoxetine (one adult, one child) have resulted in the addition of a warning in the product labeling: “Postmarketing reports indicate that STRATTERA can cause severe liver injury in rare cases. Although no evidence of liver injury was detected in clinical trials of about 6000 patients, there have been two reported cases of markedly elevated hepatic enzymes and bilirubin, in the absence of other obvious explanatory factors, out of more than 2 million patients during the first two years of postmarketing experience. In one patient, liver injury, manifested by elevated hepatic enzymes (up to 40 X upper limit of normal (ULN)) and jaundice (bilirubin up to 12 X ULN), recurred upon rechallenge, and was followed by recovery upon drug discontinuation providing evidence that STRATTERA caused the liver injury. Such reactions may occur several
months after therapy is started, but laboratory abnormalities may continue to worsen for several weeks after drug is stopped. Because of probable underreporting, it is impossible to provide an accurate estimate of the true incidence of these events. The patients described above recovered from their liver injury, and did not require a liver transplant. However, in a small percentage of patients, severe drug-related liver injury may progress to acute liver failure resulting in death or the need for a liver transplant. STRATTERA should be discontinued in patients with jaundice or laboratory evidence of liver injury, and should not be restarted. Laboratory testing to determine liver enzyme levels should be done upon the first symptom or sign of liver dysfunction (e.g., pruritus, dark urine, jaundice, right upper quadrant tenderness, or unexplained “flu-like” symptoms).\textsuperscript{112}

**Pemoline.** In 1999, an FDA review of postmarketing experience with pemoline disclosed ten reports from the U.S. of acute liver failure in children, with additional reports of liver failure in adults and in children from foreign countries. This resulted in a boxed warning stipulating that pemoline should not be used as first-line treatment for ADHD.

**Other.** A recent study by El-Zein et al. in *Cancer Letters (in press, available on web sit: http://www.sciencedirect.com/science?_ob=MImg&_imagekey=B6T54-4FH0DHK-4-1&_cdi=4992&_user=1072900&_orig=search&_coverDate=02%2F16%2F2005&_sk=999999999&view=c&wchp=dGLbVtz-zSkWb&md5=a645ef3c06f245f8b3aebf1c2bb705a&ie=/sdarticle.pdf)*, with outcomes outside the scope of this review, has raised the question of MPH’s potential cytogenetic effects. This has prompted multiple studies to assess MPH’s clastogenic potential. Results of these studies relevant to this review will be included in future updates as they become available.

**Key Question 3: Are there subgroups of patients based on demographics (age, racial groups, gender, ethnicity), other medications, or co-morbidities for which one pharmacologic treatment is more effective or associated with fewer adverse events?**

ADHD subtypes, comorbidities, and race or ethnicity were not recorded in most randomized controlled trials and observational studies. For example, only one-quarter of all studies of school-aged children reported ADHD subtype prevalence rates. Importantly, those that did record this information did not stratify or analyze data by these characteristics. While the data available from the studies that do report this information can be useful in determining the generalizability of results, the lack of attention to assessing the impact of these factors means there is no evidence on potential differences in response or adverse events.

**RACE OR ETHNICITY**

Only one-half of all studies of elementary school-aged children reported race or ethnicity among the baseline characteristics. Study populations were made up primarily of White participants, with a few exceptions. The scales used in the trials included may not perform well in all ethnic groups, or when translated into languages other than English. Since the majority of trials were performed in English speaking populations, with primarily white participants, these issues were not explored in the studies.
MPH IR. MPH IR 0.15, 0.30 and 0.50 mg/kg was studied in a placebo-controlled, crossover trial (2 weeks in each arm) of 11 Black male adolescents (mean age=13.6 years).\textsuperscript{102, 147} MPH IR had a positive effect on 75% of efficacy measures. This response rate is similar to that seen in other placebo-controlled trials of MPH IR. MPH IR was associated with significant linear elevations diastolic blood pressure among these patients.

An analysis of California Medicaid claims data suggests that mean persistence (days of treatment without any 30-day gaps) was longer for children taking MPH ER formulations (OROS and SODAS) than for those taking MPH IR regardless of ethnicity (White, Black, Hispanic)\textsuperscript{40}. This same data indicates that mean treatment durations overall (MPH OROS, SODAS, and IR) were significantly shorter for children of Black (survival time ratio (STR) 0.77; 95% CI 0.73-0.80), Hispanic (STR 0.81; 95% CI 0.78-0.84) and other ethnicities (STR 0.81; 95% CI 0.75-0.87) than for White children.

Clonidine. Clonidine 4-, 6-, and 8-mcg/kg/day was studied in a placebo-controlled crossover trial (2 weeks in each arm) of 10 children in India (80% male; mean age=7.6 years) with hyperkinetic disorder and co-occurring mental retardation.\textsuperscript{170} Clonidine was associated with dose-related improvements in 90% of all parent and clinician ratings compared to baseline. Comparisons of clonidine and placebo were not reported due to insignificant differences between placebo and baseline for any study period. Fifty percent of children developed at least mild drowsiness while taking clonidine.

Because of the small sample sizes, and limited studies of clonidine overall and of both drugs in these subgroups, indirect meta-analyses are not possible. These studies provide only limited information on the relative efficacy of MPH IR and clonidine in these groups.

GENDER

Girls typically make up only a small proportion of the total children enrolled in ADHD trials, which reflects the differential in the rates of ADHD diagnoses among the sexes.

In a study designed to assess differences in response between hyperactive boys and girls, 20 girls and 20 boys, mean age 6 years were enrolled in a randomized crossover study of placebo, MPH IR 0.15 mg/kg/day and 0.35 mg/kg/day each given for 7 to 10 days.\textsuperscript{171} Outcomes assessed were based on mother-child interactions (e.g., mother gives command, does child comply?). At baseline, there were some differences between boys and girls. The mothers of boys gave more directives during free play to boys, and during the task period, boys were less compliant with their mothers’ commands. Overall, there were no drug effects on measures during free play in any group/drug dose combination. The comparison of effects between the sexes revealed that during the task period, boys were significantly more compliant and mothers gave fewer commands and more praise comments than in the girls group. Considering the differences at baseline, this could be interpreted to indicate that the MPH IR was more effective in boys than girls in improving mother-child interactions. However, the authors do not interpret the data this way, but rather that mothers respond differently to improvements in behavior in boys than girls. These outcome measures are not used in other studies, however, and relating these findings to other studies is not possible.

A second study designed to assess differences between the response in boys and girls when taking MPH IR enrolled 24 children.\textsuperscript{172} Children were randomly assigned to placebo or MPH IR, and then crossed over to the other treatment. The randomization was done daily, with 5 to 9 days of data recorded for each condition. A number of outcome measures were used. The MANOVA analysis of results indicated a significant effect of MPH IR, but found no interaction between drug and gender.
In a study of 42 girls, analyses were primarily conducted combining data for MPH IR and DEX IR and making an indirect comparison to a study of boys conducted by the same group of researchers earlier. This report concludes that there are no striking differences between boys and girls in response to these 2 stimulants, and that both can be effective in either group.

Data from girls enrolled in 2 separate placebo-controlled trials of atomoxetine with identical protocols were analyzed post-hoc to assess the effects in this subgroup of children. These placebo-controlled trials are reported in full above. This analysis of 52 girls reported similar efficacy to that reported for the whole trial group (atomoxetine superior to placebo on most measures) but did not make a comparison of the effects in boys versus girls.

Extremely limited adverse event data was provided in these studies, and no comparison between boys and girls can be made on these measures.

**CO-MORBIDITY**

Rates of comorbidities were reported in 48% of all studies. With the exception of depression, the ranges of comorbidities reported in these trials encompass the American Academy of Pediatrics estimates on prevalence of common comorbidities: Oppositional defiant disorder=35.2 (27.2, 43.8), conduct disorder=25.7 (12.8, 41.3), anxiety disorder=25.8 (17.6, 35.3) and depressive disorder=18.2 (11.1, 26.6). One placebo-controlled trial each of atomoxetine and pemoline in adults reported results of subgroup analyses stratified by comorbidities. Atomoxetine treatment effects were not altered by the presence or absence of “psychiatric comorbidity” in a 3-week trial of 22 adults. Response to pemoline treatment was also not affected by presence or absence of “psychiatric comorbidity” in a 4-week crossover trial of 35 adults. Adults with a history of anxiety disorders tended to experience less benefit from pemoline, but with a small sample size the difference was not statistically significant. Neither trial provides evidence of comparative efficacy among subgroups of patients with comorbidities.

**Tic disorders including Tourette’s Disorder**

A concern over stimulant drugs increasing or worsening tic disorders has resulted in several trials enrolling children comorbid with ADHD and a tic disorder. One head to head and six placebo controlled trials enrolled children with tic disorders. Stimulants have been thought to worsen the severity and/or frequency of tics, hence four placebo controlled and one head to head study evaluates the effects of MPH IR in children with tics. Two additional studies examine the effects of non-stimulant medications (clonidine and guanfacine) in children with tics. Four of these studies assessed the effect of the drugs on tics using the YGTSS scale; the results are summarized in Table 15. Based on these four studies, it appears that there is no significant effect of MPH IR compared to placebo or clonidine with respect to severity of tics. A single, small study indicates that guanfacine reduced tic severity compared to placebo. Based on a variety of other measures, including tic counts by teachers and other observers, the Global Tic Rating Scale, and the Tourette Syndrome Unified Rating Scale and others, no clear evidence of MPH IR worsening tic severity or frequency could be found. Individual items were found either improved or worsened in a few instances, although a dose response for these was not found and global ratings did not find a difference compared to placebo in placebo-controlled trials. All of the placebo-controlled studies also found the MPH IR to have a positive effect on ADHD symptoms, compared to placebo.
The head-to-head study of clonidine versus MPH IR found a slight increase in tic severity between weeks 4 and 8 but, at the end of the study (16 weeks) there was no statistically significant difference between the groups.46 The proportion of children reporting a worsening of tics as an adverse effect was similar in those treated with MPH IR (20%) than those being administered clonidine alone (26%) or placebo (22%). No differences were found based on the primary outcome measure (Conners ASQ-Teacher), although individual factors indicated that clonidine was better for impulsivity and hyperactivity; while MPH IR was better for inattention. Overall, however, clonidine was significantly better than placebo on more Tourette’s outcome measures than MPH IR, and MPH IR was significantly better than placebo on more ADHD outcome measures than clonidine.

### Mental Retardation
Six randomized crossover trials of MPH IR versus placebo in children with mental retardation and ADHD (all conducted by the same group of researchers) were found.177-182 The trials were small, ranging in size from 11 to 44 children. One assessed only adverse events,178 and another enrolled preschool children.182 Most enrolled predominantly white children, with one exception.170 All participants in the Agarwal 2001 study were Indian.170 All children enrolled had mild to borderline mental retardation, as described by the eligibility criteria in each study. All of these studies had a 7 day treatment phase, and assigned patients to 0.3 and 0.6 mg/kg doses given twice daily. Additionally, one small randomized crossover trial (n = 10) of clonidine in children comorbid with mental retardation was found.170 This study differed from the MPH IR studies in that the level of mental retardation was mild in 4, moderate in 5 and severe in 1. Three clonidine doses were assigned (4, 6, or 8 mcg/kg/day) or placebo; but the randomization was to placebo or clonidine first and the order of clonidine doses was ascending. Children received each assigned treatment for 2 weeks. Tables 16 and 17 summarize the findings of these studies.

---

**Table 15. Tic Severity Ratings in trials of drugs to treat ADHD**

<table>
<thead>
<tr>
<th>Study</th>
<th>Design/Dose</th>
<th>Drug</th>
<th>YGTSS</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Placebo Controlled Studies</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Scahill 2001</td>
<td>Guanfacine</td>
<td>Placebo</td>
<td>Mean change from baseline: -4.5</td>
</tr>
<tr>
<td>Placebo</td>
<td></td>
<td></td>
<td>Mean change from baseline: 0  Effect Size: 0.67</td>
</tr>
<tr>
<td>Sverd 1992</td>
<td>MPH IR</td>
<td>Placebo</td>
<td>Global Severity score: 37.4</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>0.1 mg/kg/day: 33.0  0.3 mg/kg/day: 38.4  0.5 mg/kg/day: 35.5  ANOVA NS</td>
</tr>
<tr>
<td>Gadow 1995</td>
<td>MPH IR</td>
<td>Placebo</td>
<td>Data NR</td>
</tr>
<tr>
<td>Placebo</td>
<td></td>
<td></td>
<td>Data NR  ANOVA group x dose  NS</td>
</tr>
<tr>
<td><strong>Head to Head Study</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>The Tourette’s Study Group 2002</td>
<td>Clonidine</td>
<td>0.25mg</td>
<td>Treatment Effect Motor: 2.1, p=0.05  Vocal: 2.4, p=0.05;  OI: 6.3, p=0.007;  Total: 10.9, p=0.003</td>
</tr>
<tr>
<td></td>
<td>MPH IR 26mg</td>
<td>Placebo</td>
<td>Treatment Effect Motor: 1.3, p=NS  Vocal: 1.3, p=NS  OI: 5.8, p=0.01  Total: 9.4, p=0.01  Difference in Mean change from baseline  NS</td>
</tr>
</tbody>
</table>

---
### Table 16. Main Outcomes in Studies in Children with ADHD and Mental Retardation

<table>
<thead>
<tr>
<th>Study</th>
<th>N</th>
<th>Age</th>
<th>IQ*</th>
<th>Main Outcomes</th>
</tr>
</thead>
<tbody>
<tr>
<td>MPH IR Handen</td>
<td>12</td>
<td>NR</td>
<td>65</td>
<td>6 of 8 (75%) of items assessed during the week by the children’s teachers and 14 of 21 (67%) of items assessed during laboratory sessions showed one or both MPH IR doses superior to placebo. Teachers did not find any effect on the numbers of independent tasks completed or % correct with either dose. While teachers found both doses superior to placebo, the laboratory assessments indicated the 0.6 mg/kg dose to be superior to placebo on more measures.</td>
</tr>
<tr>
<td>MPH IR Handen</td>
<td>14</td>
<td>9 yrs</td>
<td>65</td>
<td>6 of 9 (67%) items assessed during the week by the children’s teachers and 13 of 38 (34%) of items assessed during laboratory sessions showed one or both MPH IR doses superior to placebo. Teachers rated 0.6 mg/kg better than placebo on more measures, but did not find any effect on the numbers of independent tasks completed or % correct with either dose. Conversely, the laboratory assessments indicated the 0.3 mg/kg doses to be superior to placebo on more measures. 64% classified as responders to MPH (dose groups combined, defined as 40% decrease in CTRS hyperactivity).</td>
</tr>
<tr>
<td>MPH IR Handen</td>
<td>22</td>
<td>9 yrs</td>
<td>64</td>
<td>This study only assessed free play and a “restricted academic task” (toys in a room and the adult leaves the room but the child is told not to touch the toys). 60% of each of the dependent measures showed MPH IR superior to placebo with the 0.6 mg/kg dose better on more measures than the 0.3 mg/kg dose.</td>
</tr>
<tr>
<td>MPH IR Handen</td>
<td>44</td>
<td>9 yrs</td>
<td>64</td>
<td>Overall, 66% were classified as responders (dose groups combined, defined as 50% decrease in CTRS hyperactivity). However, of these 29 children, 13 were responders only in the laboratory classroom, and 7 only in the weekday classroom. 9 (20%) were responders in both. All 5 measures by teachers, and 5 behavior measures assessed in the laboratory classroom found MPH superior to placebo. Only the Selective Reminding Test (short-term auditory learning) showed no drug effect.</td>
</tr>
<tr>
<td>MPH IR Preschool Children Handen</td>
<td>11</td>
<td>5 yrs</td>
<td>60</td>
<td>5 of 34 measures indicated 0.6 mg/kg MPH IR superior to placebo: CTRS hyperactivity, inattention-passivity, hyperactivity index; Preschool Behavior Questionnaire: hyperactive-distractible, and Play Session: intensity. No measures significantly better than placebo with 0.3 mg/kg/day. 73% classified as responders to MPH IR (dose groups combined, defined as 40% decrease in CTRS hyperactivity or PBQ hyperactive-distractible scores).</td>
</tr>
<tr>
<td>Clonidine Agarwal</td>
<td>10</td>
<td>8 yrs</td>
<td>48</td>
<td>Parent ratings (Parent Symptom Questionnaire) indicated a dose-related effect on overall behavior and hyperactive-impulsive ratings. Clinician ratings (Hillside Behavior Rating Scale) also showed clonidine to be effective, but a difference between the 6 and 8 mcg/kg dose was only apparent on a few measures. Based on the CGI-I, a dose-response was seen with 8 mcg/kg having the best result.</td>
</tr>
</tbody>
</table>

*Mean, as assessed by the Wechsler Intelligence Scale for Children – Revised
Table 17. Adverse Events in Studies in Children with ADHD and Mental Retardation

<table>
<thead>
<tr>
<th>Study</th>
<th>Adverse Events</th>
</tr>
</thead>
<tbody>
<tr>
<td>MPH IR</td>
<td></td>
</tr>
<tr>
<td>Handen 1990</td>
<td>Increased staring with MPH IR in 50%, increased drowsiness in 42%. Severe social withdrawal seen at 0.3 mg/kg doses in 1 patient.</td>
</tr>
<tr>
<td>Handen 1992</td>
<td>Not reported</td>
</tr>
<tr>
<td>Handen 1995</td>
<td>Not reported</td>
</tr>
<tr>
<td>Handen 1996</td>
<td>Not reported</td>
</tr>
<tr>
<td>Handen 1991</td>
<td>3 of 13 adverse events assessed indicated MPH IR caused fewer reports than placebo (both doses caused significantly less “high activity”, and the 0.6 mg/kg dose caused fewer episodes of irritability and anxiety.) The intensity of adverse events was assessed in 14 children, 5 of 10 adverse events assessed showed lower intensity with MPH IR than placebo (staring, irritability, anxiety, moody, and high activity).</td>
</tr>
<tr>
<td>Preschool Children</td>
<td></td>
</tr>
<tr>
<td>Handen 1999</td>
<td>5/11 had “significant adverse effects”. Social withdrawal in 6/11; severe in 4 – with the higher dose.</td>
</tr>
<tr>
<td>Clonidine</td>
<td></td>
</tr>
<tr>
<td>Agarwal 2001</td>
<td>50% had drowsiness on clonidine, resolved over 6 weeks of study in all but 2. 1 child had dry mouth and anorexia. A dose related drop in systolic blood pressure seen (3% with 4 mcg/kg, 7% with 6 mcg/kg, 9% with 8 mcg/kg).</td>
</tr>
</tbody>
</table>

Taken together, these studies indicate that MPH IR and clonidine is effective in improving some measures of ADHD symptoms. Adverse events were common, with increased staring and social withdrawal being prominent with MPH IR, and drowsiness and blood pressure changes common with clonidine. Unfortunately, these do not provide comparative evidence with other drugs.

**Autism**

A very small (n = 13) randomized crossover trial of placebo and 2 doses of MPH IR enrolled 13 children with ADHD and autism or pervasive developmental disorder (PDD) (mean age 7 years). The children’s cognitive function ranged from severely mentally retarded (23%) to normal (average IQ, 8%). Two children (15%) did not have a diagnosis of ADHD, but oppositional defiant disorder only. Overall, 54% were comorbid with oppositional defiant disorder. Additionally, one child (each) had Mosaic Downs Syndrome, Sleep-wake schedule disorder, anxiety disorder, and mixed seizure disorder and 31% had a diagnosis of PDD Not Otherwise Specified rather than autism. This study is unique in that it enrolled a higher than average proportion of black (54%) and Hispanic (15%) children. Seven children were outpatients enrolled in special education programs while 6 were inpatients or in intensive day-treatment programs. The children were randomized to MPH IR 0.3 mg/kg./day, 0.6 mg/kg./day or placebo, with the lower dose of MPH IR always preceding the higher dose. Dosing could be divided into 2 or 3 daily doses, at the choice of parents. Outcome measures assessed by teachers or program staff at the end of a 7-day dosing period included the Conners Teacher Scale, Hyperactivity Index, the IOWA CTRS, the Aberrant Behavior Checklist (ABC), and the Child Autism Rating Scale (CARS). Sixty-two percent were classified as responders to MPH IR (any dose, response = 50% reduction on the Conners Teachers Scale Hyperactivity Index). Based on final score at 7 days, both MPH doses were superior to placebo on the Conners and IOWA scales, but on the ABC scale, the 0.6 mg/kg./day dose was superior to placebo on hyperactivity ratings, and both were superior on the inappropriate speech ratings. The other 3 items did not indicate a difference from placebo for MPH IR. No difference from placebo was found the autism rating scale (CARS). An analysis of age and IQ did not reveal differences in response rates based on these factors. Eleven adverse events were reported across the three medication conditions, with “sad/unhappy/depressed” increasing with increasing doses of MPH IR and...
drowsy/dull remaining consistent across the MPH IR treatments. The other reported adverse events decreased across treatment conditions in order assigned (the effect was considered transient).

**Epilepsy**

A small \((n=30)\) randomized crossover study of children (mean age 10 years) with ADHD and epilepsy studied the effect of adding placebo or MPH IR to the child’s current anti-epilepsy regimen.\(^{183}\) The study design was unusual; children were followed for 2 months as a baseline period, then given MPH IR (open-label, uncontrolled) for 2 months however, on 2 days (one at the end of the baseline period, and one at the end of the MPH IR open label period). On these days, half of the children were given a single dose of MPH IR, and the other half placebo (non-random assignment). Outcome measures for these two days were antiepileptic drug serum levels, encephalography, and the continuous performance test (CPT). Of these, only the CPT meets criteria for this review. MPH IR was shown superior to placebo on the CPT based on speed of response and more “time on task” during the 45-minute test. The data presented on adverse events relates primarily to an observational period, although is not presented clearly. Loss of appetite was reported as an adverse event related to MPH IR that was not persistent. All others were assessed as being transient.

**Substance Abuse Disorders**

A single trial enrolled 69 adolescents (mean age 16 years) with ADHD and comorbid substance use disorder (SUD). The teens were randomized to 12 weeks of pemoline versus placebo. ADHD outcome measures were the CGI-I (investigator rated) and the Conners Hyperactivity-Impulsivity subscale of the CPRS. Outcome measures for SUD were past 30-day reports of substance use, and urine drug screens. Because conduct disorder was also thought to be a comorbidity in most patients with SUD, outcome measures for this were also included. Based on a CGI-I of 1 or 2 at endpoint, significantly more teens on pemoline were responders (32%) than those on placebo (12%, \(p<0.05\)). There was no difference between the groups on the parent rating scales, based on an ITT analysis. No differences were found on SUD or conduct disorder outcomes.

One placebo-controlled trials of MPH in adult patients with ADHD and current cocaine dependence does not provide comparative evidence.\(^{128}\) This placebo-controlled, 13-week, parallel-group trial of MPH IR reported results for 48 patients with ADHD (DSM-IV) and current cocaine dependence. Outcomes were measured using the ADHD Symptom Checklist, Global Improvement Scale (GIS), and Beck Depression Inventory. Mean participant and physician ratings at the last visit on the 7-point GIS were significantly greater in the MPH group compared to the placebo group. MPH and placebo had similar effects on ADHD symptom outcomes. Rates of adverse events were similar for MPH and placebo, with one exception. A significantly greater proportion of patients experienced insomnia or trouble sleeping while taking MPH than placebo (63% vs 33%, \(p<0.05\)).

**Limitations of this Review**

As with other types of research, it is important to recognize the limitations of this systematic review. These can be divided into those relating to generalizability of the results and those relating to methodology within the scope of this review. The generalizability of the results are limited by the scope of the key questions and inclusion criteria, and the generalizability of the
studies included. The great majority of studies included narrowly or poorly defined patient populations who met strict criteria for case definition, had few comorbidities, and used few or no concomitant medications. One concern about this group of studies is the variation in diagnostic criteria, particularly comparing studies conducted recently to those conducted in previous decades. Another concern is the handling of subtypes of ADHD in these studies. While many studies identify the proportions of patients diagnosed with various subtypes, stratification or analysis of the results based on these is lacking. Similarly, common co-morbid conditions are not well addressed by the studies. In large part, the failure to address either subtypes or co-morbidities may be due to small sample sizes involved in most studies, but these are serious short-comings that should not be ignored. The failure of these studies to assess the effect of prior medication exposure or concurrent treatment with other psychoactive medications on outcomes is another serious issue, particularly when comparing older studies where very few patients had prior exposure to newer studies where large proportions did have exposure. Minorities and the most seriously ill patients were underrepresented.

Methodological limitations of the review within the defined scope include the exclusion of studies published in languages other than English, and the lack of a specific search for unpublished studies.
## OVERALL SUMMARY

### Table 18. OVERALL SUMMARY TABLE

<table>
<thead>
<tr>
<th>Key Question 1: Benefits</th>
<th>Quality of Evidence</th>
<th>Conclusion</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>General</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Effectiveness</td>
<td>Poor, no trials found</td>
<td>No conclusions about comparative effectiveness of different pharmacotherapies for ADHD can be made</td>
</tr>
<tr>
<td><strong>Young children</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Efficacy</td>
<td>Overall: Poor</td>
<td></td>
</tr>
<tr>
<td>MPH IR</td>
<td></td>
<td>MPH IR was superior to placebo on CPRS-R efficacy outcomes</td>
</tr>
<tr>
<td><strong>Children</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Efficacy</td>
<td>Overall: Fair (individual ratings below)</td>
<td></td>
</tr>
<tr>
<td><strong>Stimulants</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>IR vs SR formulations</td>
<td>MPH IR vs MPH SR</td>
<td>Studies of MPH IR versus extended release formulations in children did not show an overall difference in efficacy.</td>
</tr>
<tr>
<td></td>
<td>(fair)</td>
<td></td>
</tr>
<tr>
<td>SR vs SR formulations</td>
<td>MPH SR vs MPH SR formulations</td>
<td>Very limited evidence suggests that Ritalin LA® was superior to Concerta® on some, but not all efficacy outcomes.</td>
</tr>
<tr>
<td></td>
<td>(poor)</td>
<td></td>
</tr>
<tr>
<td>IR vs IR</td>
<td>DEX vs MPH IR</td>
<td>The body of evidence clearly indicates no difference in efficacy between DEX and MPH IR.</td>
</tr>
<tr>
<td></td>
<td>(good)</td>
<td></td>
</tr>
<tr>
<td>Amphetamine mixture vs MPH IR</td>
<td>(fair)</td>
<td>Amphetamine mixture was superior to MPH IR on a few efficacy outcome measures in two trials, but clear evidence of superiority is lacking.</td>
</tr>
<tr>
<td>DEX IR vs DEX ER vs Amphetamine mixture</td>
<td>(poor)</td>
<td>Evidence on the comparison of DEX IR versus SR versus Amphetamine mixture may suggest that measures made in the morning show DEX IR superior to DEX SR, and afternoon measures show DEX SR superior to Amphetamine mixture.</td>
</tr>
<tr>
<td>Pemoline vs MPH IR</td>
<td>(poor)</td>
<td>Studies of Pemoline were poor quality and no conclusions can be drawn.</td>
</tr>
<tr>
<td>Modafinil (poor)</td>
<td></td>
<td>Very limited evidence from placebo-controlled trials suggests modafinil is superior to placebo on most efficacy measures.</td>
</tr>
<tr>
<td>Dexamethasone (NA)</td>
<td></td>
<td>Only incomplete evidence was found.</td>
</tr>
<tr>
<td><strong>Nonstimulants</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Atomoxetine vs MPH IR</td>
<td>(poor)</td>
<td>Limited evidence suggests a lack of a difference in efficacy compared to MPH IR.</td>
</tr>
<tr>
<td>Bupropion vs MPH IR</td>
<td>(poor)</td>
<td>Limited evidence suggests a lack of a difference in efficacy compared to MPH IR.</td>
</tr>
<tr>
<td>Clonidine vs MPH IR</td>
<td>(fair)</td>
<td>There is no clear evidence of a difference between clonidine and MPH IR among children with ADHD but no Tourette’s Disorder. Clonidine may be superior to MPH IR on tic outcome measures, with differences in ADHD outcomes in children with Tourette’s Disorder.</td>
</tr>
<tr>
<td>Guanfacine (poor)</td>
<td></td>
<td>Very limited evidence from placebo-controlled trials suggest guanfacine is superior to placebo on most efficacy measures.</td>
</tr>
<tr>
<td>Risperidone (poor)</td>
<td></td>
<td>Only post-hoc subgroup analyses of children with comorbid ADHD (placebo-controlled only).</td>
</tr>
<tr>
<td>Aripiprazole, clozapine, quetiapine, and ziprasidone (NA)</td>
<td>No evidence was found for these atypical antipsychotics.</td>
<td></td>
</tr>
<tr>
<td>Adolescents</td>
<td>Efficacy</td>
<td>Fair</td>
</tr>
<tr>
<td>-------------</td>
<td>----------</td>
<td>------</td>
</tr>
<tr>
<td>** MPH IR vs MPH OROS**</td>
<td>Functional capacity: NR</td>
<td></td>
</tr>
<tr>
<td>Short-term improvements of core ADHD symptoms: NR</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Driving performance: MPH OROS &gt; MPH IR in evening and at night</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Indirect comparisons from placebo-controlled studies of MPH IR or pemoline</strong></td>
<td>Functional capacity: NR</td>
<td></td>
</tr>
<tr>
<td>Short-term improvements of core ADHD symptoms: MPH IR and pemoline both generally efficacious; indirect comparisons not possible</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Adults</th>
<th>Efficacy</th>
<th>Fair</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Stimulants – short-term trials</strong></td>
<td>There is no evidence that any one stimulant is more effective than another</td>
<td></td>
</tr>
<tr>
<td>** DEX IR vs MPH IR (Fair)**</td>
<td>One HTH trial suggests no differences in symptom improvement</td>
<td></td>
</tr>
<tr>
<td>Indirect comparisons from placebo-controlled trials does not suggest any clear differences in response rates for DEX IR (1 trial) and MPH IR (4 trials)</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Adderall (Poor)</strong></td>
<td>Limited indirect comparisons from placebo-controlled trials suggests that response rates are similar to other stimulants</td>
<td></td>
</tr>
<tr>
<td><strong>Pemoline (Poor)</strong></td>
<td>Limited indirect comparison to stimulants from placebo-controlled trials is inconclusive</td>
<td></td>
</tr>
</tbody>
</table>

| Stimulants – longer-term observational study | |
| ** MPH IR (Poor)** | MPH IR group had lower rates of current psychiatric care than the untreated group in a 10-12 year follow-up study of males diagnosed with hyperactivity in childhood. Other outcomes were mixed. |

| Nonstimulants – short-term trials | There is no evidence that any one nonstimulant is more effective than any stimulant or other nonstimulant |
| **Atomoxetine (Poor)** | Limited indirect comparisons from placebo-controlled trials suggests that response rates are similar to stimulants and other nonstimulants. Mixed effects on quality of life compared to placebo. |
| **Bupropion SR (Fair)** | One head-to-head trial and indirect comparisons from placebo controlled trials suggests that response rates are similar to stimulants and other nonstimulants. |
| **Bupropion XL (Poor)** | Limited indirect comparisons from placebo-controlled trials suggests that response rates are similar to stimulants and other nonstimulants. |
| **Guanfacine (Poor)** | One head-to-head trial suggests no differences in positive effects on ADHD Rating Scale scores compared to DEX IR |
| **Modafinil (Poor)** | One head-to-head trial suggests no differences in response rates compared to DEX IR |
### Key Question 2: Safety

#### Quality of Evidence

**Short-term Trial Evidence**

**Young children**
- Poor – 1 placebo-controlled trial of MPH
  - Indirect comparisons cannot be made; MPH associated with higher rates of adverse events than placebo

**Children**
- Poor
  - Very few studies reported methods for assessing adverse events a priori
  - MPH IR vs MPH SR
    - There is no evidence of a difference in adverse events between IR and SR formulations.
  - MPH SR vs MPH SR formulations
    - No differences in adverse events were found.
  - DEX vs MPH IR
    - Limited evidence from short-term trials suggests that weight loss is greater with DEX than MPH IR
  - Amphetamine mixture vs MPH IR
    - Very limited evidence suggests that twice daily dosing of Amphetamine mixture led to higher rates of loss of appetite and sleep trouble.

**Adolescents**
- Poor
  - Very few studies reported methods for assessing adverse events a priori
  - MPH IR vs MPH OROS
    - NR
  - Indirect comparisons from placebo-controlled studies of MPH IR and pemoline
    - Indirect comparisons cannot be made due to heterogeneity in duration, patient characteristics, and outcome reporting

**Adults**
- Poor
  - Very few studies reported methods for assessing adverse events a priori
  - Stimulants
    - Adderall, MPH IR, and Pemoline
      - Indirect comparisons from placebo-controlled trials suggest all are associated with higher rates of insomnia, appetite loss and withdrawal due to adverse events than placebo
    - DEX IR and MPH SR
      - Indirect comparisons cannot be made.
  - Nonstimulants
    - Atomoxetine
      - Indirect comparisons suggest that atomoxetine is associated with rates of insomnia, appetite loss and withdrawals due to adverse events similar to stimulants
    - Bupropion SR vs MPH IR
      - No differences in adverse events in one head-to-head trial
    - Bupropion XL
      - Associated with insignificant increases in insomnia and withdrawals due to adverse events. Indirect comparisons to stimulants are inconclusive.
    - Guanfacine vs DEX IR
      - No difference in adverse events in one head-to-head trial
    - Modafinil vs DEX IR
      - No differences in adverse events in one head-to-head trial
<table>
<thead>
<tr>
<th>Long-Term Safety – Observational Studies</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mixed populations, primarily children</td>
</tr>
<tr>
<td><strong>Height</strong></td>
</tr>
<tr>
<td>• DEX vs MPH IR: Mixed findings. DEX=MPH in 6-year height increases in one study; DEX&gt;MPH in 2-year height decreases in the other</td>
</tr>
<tr>
<td>• MPH IR vs unmedicated controls: No significant differences in two studies</td>
</tr>
<tr>
<td>• MPH IR in uncontrolled studies: Inconsistent effects across four studies</td>
</tr>
<tr>
<td>• Atomoxetine: Uncontrolled studies suggest that height changes are similar to those reported with MPH IR, and are also transient</td>
</tr>
<tr>
<td><strong>Weight</strong></td>
</tr>
<tr>
<td>• DEX vs MPH: Three studies consistently suggest that DEX&gt;MPH in weight gain suppression in the first 1-2 years. The longest-term (5 years) of these studies also reported that DEX=MPH in exceeding weight gain expectations at final follow-up. These findings are weakened by methodological flaws, however.</td>
</tr>
<tr>
<td>• MPH IR in other comparative (imipramine and unmedicated hyperactives or healthy controls) and noncomparative studies: Evidence does not support an indisputable relationship between MPH and weight gain suppression</td>
</tr>
<tr>
<td>• MPH OROS and atomoxetine: Evidence from noncomparative studies (one each) doesn’t suggest weight gain suppression effects</td>
</tr>
<tr>
<td>• Atomoxetine: Uncontrolled studies suggest that weight changes are similar to those reported with MPH IR, and are also transient.</td>
</tr>
<tr>
<td><strong>Tics, seizures, cardiovascular adverse events, injuries, and attempted suicide</strong></td>
</tr>
<tr>
<td><strong>Drugs with warnings or removal from market</strong></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>Key Question 3: Subgroups</td>
</tr>
<tr>
<td>--------------------------</td>
</tr>
<tr>
<td><strong>Children</strong></td>
</tr>
<tr>
<td>ADHD Subtypes or Severity</td>
</tr>
<tr>
<td>Race / Ethnicity</td>
</tr>
<tr>
<td>Age</td>
</tr>
<tr>
<td>Gender</td>
</tr>
<tr>
<td>Common Co-morbidities</td>
</tr>
<tr>
<td>Tic Disorders</td>
</tr>
<tr>
<td>Mental Retardation</td>
</tr>
<tr>
<td>Autism</td>
</tr>
<tr>
<td>Epilepsy</td>
</tr>
<tr>
<td>Substance Use Disorder</td>
</tr>
<tr>
<td>Adults</td>
</tr>
</tbody>
</table>
REFERENCES

1. NIH. NIH Consensus Statement: Diagnosis and treatment of attention deficit hyperactivity disorder. Accessed 2, 16.


116. Paternite CE, Loney J, Salisbury H, Whaley MA. Childhood inattention-overactivity, aggression, and stimulant medication history as predictors of young


Figure 1. Pharmacologic Treatments in ADHD: Drug Class Review Flow Diagram

2,287 citations

1,773 excluded at title/abstract level

514 retrieved for full-text evaluation

334 excluded at full-text level
Reasons for exclusion include not English language, wrong outcome, drug not included, population not included, wrong publication type, wrong study design, insufficient duration

180 publications included
- 40 head-to-head trials
- 67 placebo-controlled trials
- 26 observational studies
- 5 active-controlled trials
- 4 multimodal trials
- 4 systematic review
- 34 background
Appendix A. Scales Used to Assess Efficacy and Adverse Events

The following narrative briefly describes the most commonly used assessment scales and summarizes methods of scoring and validation.

**ADHD Behavior Checklist/ADHD Rating Scale** evaluates inattentive and hyperactive-impulsive symptoms is based on DSM criteria for diagnosing ADHD. DSM-III uses a 14-item checklist while DSM-IV updated it to an 18-item checklist with two nine-item subscales. Items are rated for severity from zero to three according to how often the symptoms are present (0=never/rarely, 1=sometimes, 2=often, and 3=very often). The maximum scores are 42 points and 54 points for DSM-III and DSM-IV respectively. The test-retest reliability was demonstrated. The intraclass correlation coefficient was .90s (p<0.001). The content validity and construct validity were proved as well. The checklist has established validity, reliability, and age-matched cut-off values (Achenbach & Edelbrock, 1983; DuPaul, 1998).

**Barkley’s Attention Deficit Hyperactivity Disorder Checklist and Scale** is a self-report rating system that measures the occurrence of symptoms. The range of the scale is 0=never or rarely, 1=sometimes, 2=often, and 3=very often. The checklist is used a measurement to define symptoms of the disorder. No reliability or validity information available (Barkley & Murphy, 1998).

**Barratt Impulsiveness Scale (BIS-10)** is a 34-item scale that covers three types of impulsiveness: motor, cognitive, and non-planning. It consists of a four-point scale ranging (“rarely/never”, “occasionally”, “often”, and “almost always/always”). These three factors are considered reliable under a study with an alpha coefficient range from 0.89 to 0.92. No validity information available (Barratt, 1985).

**Brown ADD scale** is a 40-item self report scale for assessing the executive function aspects associated with ADHD. The scale has been proved with good internal consistency and good test-retest reliability. The total score ranges from 0 to 120: patients with score >55 = highly probable ADHD; score 40-54 = 'probable' ADHD; score <40 = 'possible' ADHD (Brown, 1996).

**Child Behavior Checklist (CBCL)** originally had three axes, the parent report form, teacher report form, and self-report form for children over 11 years of age (Achenbach & Edelbrock, 1981). But it had been added to have two more axes, which are cognitive assessment and physical assessment from observations and interviews. It was demonstrated to have highly reliability and validity through various studies (Achenbach, 1999).

**Children’s Global Assessment Scale (CGAS)** is an adaptation of the Global Assessment Scale (GAS). This scale is designed to measure the lowest level of functioning during a specific time period for children aged 4 to 16. Children are rated on a scale of 1 (needs constant supervision) to 100 (superior functioning) with anchor points in between. Scores above 70 indicate normal function. The CGAS has demonstrated discriminate validity (P=.001) in detecting the level of impairment between inpatients and outpatients. The CGAS has also demonstrated concurrent validity with the Conners ten-item Abbreviated Parent Checklist; the correlation was –0.25 (P > .05, df=17) when used in outpatients (Shaffer et al., 1983).
Children’s Psychiatric Rating Scale (CPRS) is a comprehensive, 63-item scale that aims to assess a broad spectrum of psychopathology for children up to age 15. Therefore, items on the CPRS will have varying degrees of relevance when used in a specific diagnostic group. Each item is rated from one (not present) to seven (extremely severe). But unfortunately, we can’t find any information about the reliability and validity of the scale (Fish 1985).

Continuous Performance Test (CPT) is a monitoring task in which subjects are given a series of visual or auditory stimuli and are asked to press a button when certain, infrequent target stimuli appear. There is no standardized version. There is usually a “low-level” version and a more sophisticated version where the stimulus may or may not be a target depending on what precedes it in the series. (Conners, Erhardt, & Sparrow, 1999; Epstein, Conners, Sitarenios, & Erhardt, 1998; Halperin, Greenblatt, Sharma, & Schwartz, 1991; Nuechterlein, 1983; Rosvold, Mirsky, Sarason, Bransome Jr, & Beck, 1956)

Clinical Global Impression Scale (CGI) is used in both children and adults and consists of three global scales for rating mental illness. The first two items (severity of illness and global improvement) are rated on a seven-point scale (1 = very much improved, 7 = very much worse). The third item (efficacy index) uses a matrix to rate the effectiveness of therapy in relation to adverse reactions (Guy, 1976).

Conners’ Abbreviated Questionnaires (ASQ-P) is an abbreviated version of the CPRS. It contains 10 items only, and is known as the Hyperactivity Index. The intercorrelation of ASQ–P and CPRS-R was high as .87 in the hyperactive factor that demonstrated the ASQ-T’s ability to identify children’s hyperactive behaviors (Goyette, Conners, & Ulrich, 1978). Parents rate their child’s symptoms from zero to three (0=not at all present, 1=just a little present, 2=pretty much present, 3=very much present), which yields a range of possible total scores between 0 and 30.

Conners’ Abbreviated Questionnaires (ASQ-T) is an Abbreviated version of the CTRS. It contains 10 items only, and is known as the Hyperactivity Index. The intercorrelation of ASQ –T and CTRS-R was high from .79-.90 that demonstrated the ASQ-T’s ability to identify children’s problem behaviors (Goyette, Conners, & Ulrich, 1978)

Conners’ Parent Rating Scale (CPRS) is a 93-item parent rating scale to evaluate children’s psychiatric symptoms. It is the original version of the CPRS. Parents rate their child’s symptoms from one to four (1=not at all present, 2=just a little present, 3=pretty much present, 4=very much present) (Conners, Erhardt, & Sparrow, 1999).

The 48-item Conners’ Parent Rating Scale – Revised (CPRS-R) is a revised version of the 93-item Conners’ Parent Rating Scale and includes norms down to age three. Parents rate their child’s symptoms from zero to three (0=not at all present, 1=just a little present, 2=pretty much present, 3=very much present). (Goyette, Conners, & Ulrich, 1978)
**Conners’ Teacher Rating Scale (CTRS)** is a 39-item teacher rating scale teachers to evaluate children’s symptoms and behaviors before and after medication. The four-points scale (1-not at all, 2-just a little, 3-quite a bit, and 4-very much) was rated. Factor analysis was used to prove the stability of the scale. It is highly sensitive to drug effectiveness (Conners, Erhardt, & Sparrow, 1999). Teachers rate their child’s symptoms from zero to three (0=not at all present, 1=just a little present, 2=pretty much present, 3=very much present), which yields a range of possible total scores between 0 and 30.

The 28-item **Conners’ Teacher Rating Scale – Revised (CTRS-R)** is a revised version of the 48-item Conners’ Teacher Rating Scale and includes norms down to age three. Teachers rate their child’s symptoms from zero to three (0=not at all present, 1=just a little present, 2=pretty much present, 3=very much present). (Goyette, Conners, & Ulrich, 1978)

**Copeland Symptom Checklist for Adult Attention Deficit Disorder**, an 8-category, 63-item checklist with each item rated on a severity scale from 0 (symptoms not present) through 4 (very much present). It contents the information about cognitive, emotional and social symptoms. Its validity and reliability have been established, but we were unsuccessful in retrieving the original source, “Copeland Symptom Checklist for Adult Attention Deficit Disorders”(Copeland, 1989).

**Global Assessment Scale (GAS)** is a single rating scale for assessing the overall functioning of a patient. The scale values range from 1 to 100, with 1 being the hypothetical sickest person and 100 being the hypothetical healthiest person. There are ten equal intervals ranging from 1-10, 11-20, 21-30 and so on up until 91-100; if a patient falls in the upper two intervals, it is considered “positive mental health.” A patient is rated based on observing his behavior during the preceding week and comparing it to the current time period, and adjustments are made to base on specific characteristics defined in each interval. The GAS is found to have good reliability based on five studies with an intra-class correlation coefficient range of 0.61 to 0.95 and an associated standard error of measurement range of 5.0 to 8.0 units. Strong concurrent validity was proved as well (Endicott, Spitzer, Fleiss, & Cohen, 1976).

**“How I Feel” Questionnaire**, a 28-item scale, is an adaptation of the van Kammen-Murphy Mood Scale, which has been proved to be sensitive to the effects of amphetamine. It uses 4-point scale: 0= “not at all”; 1= ”a little”; 2= “some”; 3= “a lot”. No reliability or validity information is available (Rapoport et al., 1980).

**Inattention/Overactivity With Aggression Conners’ Teacher Rating Scale (IOWA CTRS)** is revised from the 39-item Conner’s Teacher scale. 10 items were devised to determine Inattention-Overactivity (IO) and aggression (A) behaviors. Teachers rate their child’s symptoms from zero to three (0=not at all, 1=just a little, 2=pretty much, 3=very much). Coefficient alpha was tested as .89 for the IO scale and .86 for the A scale. They only tested the sensitivity and specificity scores of the IO scale, and the scores depend on the screen score being rated. Therefore, it recommended the use of an IO scale for at least 11 points for research purpose, and 7 points for clinical purpose (Loney & Milich, 1982). The differential validity of IO and A factors had been tested as well (Atkins, Pelham, & Licht, 1989).
Physician’s Global Rating Scale is a seven-point rating of the overall functioning of a patient. The physician rates the patient improvement on a scale from –3 to +3. The number measures the change seen in the patient (-3=marked worsening, -2=moderate worsening, -1=slight worsening, 0=no change, +1=mild improvement, +2=moderate improvement, +3=marked improvement). No validity or reliability information is available (Wender, Reimherr, Wood, & Ward, 1985).

Physician’s Target Symptom Scale is a four-point rating scale, ranging from 0 to 3 (0=not at all, 1=mild, 2=moderate, 3=marked). It measures specific symptoms of attention deficit disorder: conduct disorder (CD), disorganization, depression, temper, short attention span, and hyperactivity. No validity or reliability information is available (Wender, Reimherr, Wood, & Ward, 1985).

SCL-90 Rating Scale is a self-report clinical rating scale. It uses a 90-item checklist that covers nine symptom constructs, and three global indices of pathology. It consists of a five-point scale that measures the amount of distress a patient has felt to identify symptomatic behavior of psychiatric outpatients: 0=not at all, 1=a little bit, 2=moderately, 3=quite a bit, 4=extremely. There is evidence of strong convergent validity when compared to MMPI. No reliability information is available (Derogatis, Lipman, & Covi, 1973; Derogatis, Rickels, & Rock, 1976).

Sheehan Disability Scale (SDS), a three-item instrument for assessing psychiatric impairment in occupational, social and family functioning, each rated from 0 to 10 (0-3: mild impairment; 4-6: moderate impairment; 7-10: severe impairment). Internal consistency reliability was demonstrated with the coefficient alpha was 0.89 for three-item scale. Reliability of each item ranged from 0.67 for work impairment to 0.77 for family impairment and 0.81 for social impairment. The construct validity was proved as well (Leon, Olfson, Portera, Farber, & Sheehan, 1997).

Swanson, Kotlin, Agler, M-Flynn and Pelham (SKAMP) scale is a 15-item scale. Ten items describe typical behaviors in a classroom setting and other five items were used for recording specific behavior (Swanson, 1992). Items are rated on a 7-point impairment scale (none, slight, mild, moderate, severe, very severe, and maximal). The reliabilities were from .70 to .78 for the SKAMP Attention ratings, and were from .63 to .73 for the SKAMP Deportment ratings. The concurrent validity was established by calculating correlations with Conners and the IOWA Conners Rating scale (Wigal, Gupta, Guinta, & Swanson, 1998).

Wender Utah Rating Scale (WURS) is a 61-item scale for adults to evaluate childhood behavior. It has been demonstrated to be sensitive in identifying childhood attention deficit hyperactivity disorder. It is rated on the five-point scale: 'not at all or slightly', 'mildly', 'moderately', 'quite a bit', and 'very much'. A subset of 25 of the items successfully identified 86% of patients diagnosed with ADHD and 99% of the normal, control individuals (Ward, Wender, & Reimherr, 1993). The test-retest reliability was proved with Cronbach alpha ranged from .69 to .90. The validity was demonstrates as well with factor analysis (Rossini & O'Connor, 1995; Stein et al., 1995).
Wechsler Intelligence Scale for Children, 3rd edition (WISC-III) is an instrument assessing the intellectual ability of children aged 6 to 16 years. It consists of different measures to estimate individual’s intellectual abilities. Each subtest is derived from four factors, verbal comprehension, perceptual organization, freedom from distractibility and processing speed. The reliability coefficients of the subscales are from .69-.96. Besides, it has been demonstrated in construct validity and internal validity (Wechsler, 1991).

Werry-Quay Direct Observational System assesses behaviors including out-of-seat; physical contact or disturbing others; audible noise; ninety-degree turn, seated; inappropriate vocalizations; other deviant behaviors; and daydreaming. Retrieval of reliability and validity findings (Werry & Quay, 1969) are pending and will be addressed in the updated report.

Reference


## Appendix B. Search Strategies

Database: EBM Reviews - Cochrane Central Register of Controlled Trials <2nd Quarter 2004>

Search Strategy:

<table>
<thead>
<tr>
<th>1</th>
<th>Methylphenidate.mp. (769)</th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td>concerta.mp. (7)</td>
</tr>
<tr>
<td>3</td>
<td>metadate.mp. (2)</td>
</tr>
<tr>
<td>4</td>
<td>methylin.mp. (0)</td>
</tr>
<tr>
<td>5</td>
<td>ritalin.mp. (80)</td>
</tr>
<tr>
<td>6</td>
<td>dexamethylphenidate.mp. (0)</td>
</tr>
<tr>
<td>7</td>
<td>focalin.mp. (2)</td>
</tr>
<tr>
<td>8</td>
<td>pemoline.mp. (78)</td>
</tr>
<tr>
<td>9</td>
<td>cylert.mp. (9)</td>
</tr>
<tr>
<td>10</td>
<td>amphetamine$.mp. (647)</td>
</tr>
<tr>
<td>11</td>
<td>adderall.mp. (21)</td>
</tr>
<tr>
<td>12</td>
<td>dextroamphetamine.mp. (406)</td>
</tr>
<tr>
<td>13</td>
<td>Dexedrine.mp. (14)</td>
</tr>
<tr>
<td>14</td>
<td>1 or 2 or 3 or 4 or 5 or 6 or 7 or 8 or 9 or 10 or 11 or 12 or 13 (1561)</td>
</tr>
<tr>
<td>15</td>
<td>Atomoxetine.mp. (12)</td>
</tr>
<tr>
<td>16</td>
<td>Strattera.mp. (0)</td>
</tr>
<tr>
<td>17</td>
<td>wellbutrin.mp. (13)</td>
</tr>
<tr>
<td>18</td>
<td>bupropion.mp. (287)</td>
</tr>
<tr>
<td>19</td>
<td>modafinil.mp. (72)</td>
</tr>
<tr>
<td>20</td>
<td>provigil.mp. (1)</td>
</tr>
<tr>
<td>21</td>
<td>clonidine.mp. (1917)</td>
</tr>
<tr>
<td>22</td>
<td>catapres.mp. (18)</td>
</tr>
<tr>
<td>23</td>
<td>guanfacine.mp. (95)</td>
</tr>
<tr>
<td>24</td>
<td>tenex.mp. (0)</td>
</tr>
<tr>
<td>25</td>
<td>aripiprazole.mp. (54)</td>
</tr>
<tr>
<td>26</td>
<td>clozapine.mp. (663)</td>
</tr>
<tr>
<td>27</td>
<td>clozaril.mp. (7)</td>
</tr>
<tr>
<td>28</td>
<td>olanzapine.mp. (935)</td>
</tr>
<tr>
<td>29</td>
<td>zyprexa.mp. (3)</td>
</tr>
<tr>
<td>30</td>
<td>quetiapine.mp. (144)</td>
</tr>
<tr>
<td>31</td>
<td>seroquel.mp. (97)</td>
</tr>
<tr>
<td>32</td>
<td>risperidone.mp. (958)</td>
</tr>
<tr>
<td>33</td>
<td>risperdal.mp. (7)</td>
</tr>
<tr>
<td>34</td>
<td>ziprasidone.mp. (206)</td>
</tr>
<tr>
<td>35</td>
<td>geodon.mp. (0)</td>
</tr>
<tr>
<td>36</td>
<td>15 or 16 or 17 or 18 or 19 or 20 or 21 or 22 or 23 or 24 or 25 or 26 or 27 or 28 or 29 or 30 or 31 or 32 or 33 or 34 or 35 (4755)</td>
</tr>
<tr>
<td>37</td>
<td>Attention Deficit Disorder with Hyperactivity/ (548)</td>
</tr>
<tr>
<td>38</td>
<td>attention deficit disorder/ (548)</td>
</tr>
<tr>
<td>39</td>
<td>attention deficit.mp. (767)</td>
</tr>
<tr>
<td>40</td>
<td>adhd.mp. (389)</td>
</tr>
</tbody>
</table>
Database: Ovid MEDLINE(R) <1966 to June Week 2 2004>
Search Strategy:

1  Methylphenidate.mp. (3391)
2  concerta.mp. (19)
3  metadate.mp. (8)
4  methylin.mp. (2)
5  ritalin.mp. (310)
6  dexamethylphenidate.mp. (9)
7  focalin.mp. (4)
8  pemoline.mp. (469)
9  cyler.mp. (23)
10 amphetamine$.mp. (19249)
11 adderall.mp. (52)
12 dextroamphetamine.mp. (5587)
13 Dexedrine.mp. (36)
14 1 or 2 or 3 or 4 or 5 or 6 or 7 or 8 or 9 or 10 or 11 or 12 or 13 (23723)
15 Atomoxetine.mp. (48)
16 Strattera.mp. (5)
17 wellbutrin.mp. (20)
18 bupropion.mp. (1243)
19 modafinil.mp. (256)
20 provigil.mp. (8)
21 clonidine.mp. (13960)
22 catapres.mp. (87)
23 guanfacine.mp. (562)
24 tenex.mp. (4)
25 aripiprazole.mp. (98)
26 clozapine.mp. (5713)
27 clozaril.mp. (45)
28 olanzapine.mp. (2143)
29 zyprexa.mp. (19)
30 quetiapine.mp. (1090)
31 seroquel.mp. (129)
32 risperidone.mp. (2725)
33 risperdal.mp. (16)
34 ziprasidone.mp. (348)
35 geodon.mp. (5)
36 15 or 16 or 17 or 18 or 19 or 20 or 21 or 22 or 23 or 24 or 25 or 26 or 27 or 28 or 29 or 30
or 31 or 32 or 33 or 34 or 35 (24748)
37 Attention Deficit Disorder with Hyperactivity/ (7936)
38 attention deficit.mp. (4965)
39 adhd.mp. (3052)
40 37 or 38 or 39 (9040)
41 14 or 36 (47696)
42 40 and 41 (1900)
43 limit 42 to (controlled clinical trial or meta analysis or randomized controlled trial) (429)
44 (systemat$ adj5 review$).mp. (6912)
45 Randomized Controlled Trials/ (32564)
46 cohort$.mp. (88689)
47 44 or 45 or 46 (125108)
48 42 and 47 (68)
49 43 or 48 (478)
50 adverse effect$.mp. or ae.fs. (837033)
51 poisoning.mp. or po.fs. (63877)
52 toxicity.mp. or to.fs. (265791)
53 50 or 51 or 52 (1110936)
54 41 and 53 (10260)
55 limit 54 to (controlled clinical trial or meta analysis or randomized controlled trial) (1128)
56 47 and 54 (237)
57 55 or 56 (1312)
58 49 or 57 (1649)
59 limit 58 to human (1645)
60 limit 59 to english language (1572)
61 limit 59 to abstracts (1489)
62 60 or 61 (1630)
63 from 62 keep 1-1630 (1630)

Database: Ovid MEDLINE(R) <1966 to June Week 2 2004>
Search Strategy:
-----------------------------------------------------------------------------------
1 Methylphenidate.mp. (3391)
2 concerta.mp. (19)
3 metadate.mp. (8)
4 methylin.mp. (2)
5 ritalin.mp. (310)
6 dexamethylphenidate.mp. (9)
7 focalin.mp. (4)
8 pemoline.mp. (469)
9 cylert.mp. (23)
10 amphetamine$.mp. (19249)
11 adderall.mp. (52)
dextroamphetamine.mp. (5587)  
Dexedrine.mp. (36)  
1 or 2 or 3 or 4 or 5 or 6 or 7 or 8 or 9 or 10 or 11 or 12 or 13 (23723)  
Atomoxetine.mp. (48)  
Strattera.mp. (5)  
wellbutrin.mp. (20)  
bupropion.mp. (1243)  
modafinil.mp. (256)  
provigil.mp. (8)  
clonidine.mp. (13960)  
catapres.mp. (87)  
guanfacine.mp. (562)  
tenex.mp. (4)  
aripiprazole.mp. (98)  
clozapine.mp. (5713)  
clozaril.mp. (45)  
olanzapine.mp. (2143)  
zyprexa.mp. (19)  
quetiapine.mp. (1090)  
seroquel.mp. (129)  
risperidone.mp. (2725)  
risperdal.mp. (16)  
ziprasidone.mp. (348)  
geodon.mp. (5)  
15 or 16 or 17 or 18 or 19 or 20 or 21 or 22 or 23 or 24 or 25 or 26 or 27 or 28 or 29 or 30 or 31 or 32 or 33 or 34 or 35 (24748)  
Attention Deficit Disorder with Hyperactivity/ (7936)  
attention deficit.mp. (4965)  
adhd.mp. (3052)  
37 or 38 or 39 (9040)  
14 or 36 (47696)  
40 and 41 (1900)  
limit 42 to (controlled clinical trial or meta analysis or randomized controlled trial) (429)  
(systemat$ adj5 review$).mp. (6912)  
Randomized Controlled Trials/ (32564)  
cohort$.mp. (88689)  
44 or 45 or 46 (125108)  
42 and 47 (68)  
43 or 48 (478)  
adverse effect$.mp. or ae.fs. (837033)  
poisoning.mp. or po.fs. (63877)  
toxicity.mp. or to.fs. (265791)  
50 or 51 or 52 (1110936)  
41 and 53 (10260)  
limit 54 to (controlled clinical trial or meta analysis or randomized controlled trial) (1128)  
47 and 54 (237)
Database: PsycINFO <1974 to May Week 5 2004>
Search Strategy:

1  Methylphenidate.mp. (1852)
2  concerta.mp. (4)
3  metadate.mp. (2)
4  methylin.mp. (0)
5  ritalin.mp. (248)
6  dexamethylphenidate.mp. (2)
7  focalin.mp. (1)
8  pemoline.mp. (147)
9  cylert.mp. (17)
10  amphetamine$.mp. (5905)
11  adderall.mp. (40)
12  dextroamphetamine.mp. (1980)
13  Dexedrine.mp. (29)
14  1 or 2 or 3 or 4 or 5 or 6 or 7 or 8 or 9 or 10 or 11 or 12 or 13 (8060)
15  Atomoxetine.mp. (27)
16  Strattera.mp. (0)
17  wellbutrin.mp. (10)
18  bupropion.mp. (575)
19  modafinil.mp. (95)
20  provigil.mp. (2)
21  clonidine.mp. (1577)
22  catapres.mp. (5)
23  guanfacine.mp. (82)
24  tenex.mp. (0)
25  aripiprazole.mp. (35)
26  clozapine.mp. (3467)
27  clozaril.mp. (28)
28  olanzapine.mp. (1542)
29  zyprexa.mp. (6)
30  quetiapine.mp. (500)
31  seroquel.mp. (56)
32  risperidone.mp. (2020)
Embase

((('methylphenidate'/exp OR 'concerta'/exp OR metadata OR methylin OR 'ritalin'/exp OR 'dexamphetamine'/exp OR 'focalin'/exp OR 'pemoline'/exp OR 'cyrlert'/exp OR 'amphetamine' OR 'adderall'/exp OR 'dextroamphetamine'/exp OR 'dexdrenaline'/exp AND [embase]/lim) OR ('atomoxetine'/exp OR 'strattera'/exp OR 'wellbutrin'/exp OR 'buproprion'/exp OR 'modafinil'/exp OR 'provigil'/exp OR 'clonidine'/exp OR 'catapres'/exp OR 'guanfacine'/exp OR 'tenex'/exp OR 'aripiprazole'/exp OR 'clozapine'/exp OR 'clozaril'/exp OR 'olanzapine'/exp OR 'zyprexa'/exp OR 'quetiapine'/exp OR 'seroquel'/exp OR 'risperidone'/exp OR 'risperdal'/exp OR 'ziprasidone'/exp OR 'geodon'/exp AND [embase]/lim))
AND
('attention deficit disorder'/exp OR adhd OR 'attention deficit'/exp AND [embase]/lim))
AND
([meta analysis]/lim OR [randomized controlled trial]/lim)
AND
[embase]/lim
Appendix C. Quality assessment methods for drug class reviews for the Drug Effectiveness Review Project

The purpose of this document is to outline the methods used by the Oregon Evidence-based Practice Center (EPC), based at Oregon Health & Science University, and any subcontracting EPCs, in producing drug class reviews for the Drug Effectiveness Review Project. The methods outlined in this document ensure that the products created in this process are methodologically sound, scientifically defensible, reproducible, and well documented. This document has been adapted from the Procedure Manual developed by the Methods Work Group of the United States Preventive Services Task Force (version 1.9, September 2001), with additional material from the NHS Centre for Reviews and Dissemination (CRD) report on Undertaking Systematic Reviews of Research on Effectiveness: CRD’s Guidance for Carrying Out or Commissioning Reviews (2nd edition, 2001) and “The Database of Abstracts of Reviews of Effects (DARE)” in Effectiveness Matters, vol. 6, issue 2, December 2002, published by the CRD.

All studies or systematic reviews that are included are assessed for quality, and assigned a rating of “good”, “fair” or “poor”. Studies that have a fatal flaw in one or more criteria are rated poor quality; studies which meet all criteria, are rated good quality; the remainder are rated fair quality. As the “fair quality” category is broad, studies with this rating vary in their strengths and weaknesses: the results of some fair quality studies are likely to be valid, while others are only probably valid. A “poor quality” trial is not valid—the results are at least as likely to reflect flaws in the study design as the true difference between the compared drugs.

For Controlled Trials:

Assessment of Internal Validity

1. Was the assignment to the treatment groups really random?
   Adequate approaches to sequence generation:
   - Computer-generated random numbers
   - Random numbers tables
   Inferior approaches to sequence generation:
   - Use of alternation, case record numbers, birth dates or week days
   - Not reported

2. Was the treatment allocation concealed?
   Adequate approaches to concealment of randomization:
   - Centralized or pharmacy-controlled randomization
   - Serially-numbered identical containers
   - On-site computer based system with a randomization sequence that is not readable until allocation
   - Other approaches sequence to clinicians and patients
   Inferior approaches to concealment of randomization:
   - Use of alternation, case record numbers, birth dates or week days
Open random numbers lists
Serially numbered envelopes (even sealed opaque envelopes can be subject to manipulation)
Not reported

3. Were the groups similar at baseline in terms of prognostic factors?

4. Were the eligibility criteria specified?

5. Were outcome assessors blinded to the treatment allocation?

6. Was the care provider blinded?

7. Was the patient kept unaware of the treatment received?

8. Did the article include an intention-to-treat analysis, or provide the data needed to calculate it (i.e., number assigned to each group, number of subjects who finished in each group, and their results)?

9. Did the study maintain comparable groups?

10. Did the article report attrition, crossovers, adherence, and contamination?

11. Is there important differential loss to followup or overall high loss to followup? (give numbers in each group)

Assessment of External Validity (Generalizability)

1. How similar is the population to the population to whom the intervention would be applied?

2. How many patients were recruited?

3. What were the exclusion criteria for recruitment? (Give numbers excluded at each step)

4. What was the funding source and role of funder in the study?

5. Did the control group receive the standard of care?

6. What was the length of followup? (Give numbers at each stage of attrition.)
**For Studies Reporting Complications/Adverse Effects**

**Assessment of Internal Validity**

1. Was the selection of patients for inclusion non-biased (Was any group of patients systematically excluded)?

2. Is there important differential loss to followup or overall high loss to followup? (Give numbers in each group.)

3. Were the events investigated specified and defined?

4. Was there a clear description of the techniques used to identify the events?

5. Was there non-biased and accurate ascertainment of events (independent ascertainers; validation of ascertainment technique)?

6. Were potential confounding variables and risk factors identified and examined using acceptable statistical techniques?

7. Did the duration of followup correlate to reasonable timing for investigated events? (Does it meet the stated threshold?)

**Assessment of External Validity**

1. Was the description of the population adequate?

2. How similar is the population to the population to whom the intervention would be applied?

3. How many patients were recruited?

4. What were the exclusion criteria for recruitment? (Give numbers excluded at each step)

5. What was the funding source and role of funder in the study?

**Systematic Reviews:**

1. Is there a clear review question and inclusion/exclusion criteria reported relating to the primary studies?

A good quality review should focus on a well-defined question or set of questions, which ideally will refer to the inclusion/exclusion criteria by which decisions are made on whether to include or exclude primary studies. The criteria should relate to the four components of study design, indications (patient populations), interventions (drugs), and outcomes of interest. In addition, details should be reported relating to the process of decision-making,
i.e., how many reviewers were involved, whether the studies were examined independently, and how disagreements between reviewers were resolved.

2. Is there evidence of a substantial effort to search for all relevant research?

This is usually the case if details of electronic database searches and other identification strategies are given. Ideally, details of the search terms used, date and language restrictions should be presented. In addition, descriptions of hand-searching, attempts to identify unpublished material, and any contact with authors, industry, and research institutes should be provided. The appropriateness of the database(s) searched by the authors should also be considered, e.g. if MEDLINE is searched for a review looking at health education, then it is unlikely that all relevant studies will have been located.

3. Is the validity of included studies adequately assessed?

A systematic assessment of the quality of primary studies should include an explanation of the criteria used (e.g., method of randomization, whether outcome assessment was blinded, whether analysis was on an intention-to-treat basis). Authors may use either a published checklist or scale, or one that they have designed specifically for their review. Again, the process relating to the assessment should be explained (i.e. how many reviewers involved, whether the assessment was independent, and how discrepancies between reviewers were resolved).

4. Is sufficient detail of the individual studies presented?

The review should demonstrate that the studies included are suitable to answer the question posed and that a judgement on the appropriateness of the authors' conclusions can be made. If a paper includes a table giving information on the design and results of the individual studies, or includes a narrative description of the studies within the text, this criterion is usually fulfilled. If relevant, the tables or text should include information on study design, sample size in each study group, patient characteristics, description of interventions, settings, outcome measures, follow-up, drop-out rate (withdrawals), effectiveness results and adverse events.

5. Are the primary studies summarized appropriately?

The authors should attempt to synthesize the results from individual studies. In all cases, there should be a narrative summary of results, which may or may not be accompanied by a quantitative summary (meta-analysis).

For reviews that use a meta-analysis, heterogeneity between studies should be assessed using statistical techniques. If heterogeneity is present, the possible reasons (including chance) should be investigated. In addition, the individual evaluations should be weighted in some way (e.g., according to sample size, or inverse of the variance) so that studies that are considered to provide the most reliable data have greater impact on the summary statistic.
Appendix D. Excluded Studies


160. McKeage K, Scott LJ. SLI-381 (Adderall XR(registered trademark)). *CNS Drugs.* 2003;17(9):669-675.


176. Pearson DA, Santos CW, Roache JD, et al. Treatment effects of methylphenidate on behavioral adjustment in children with mental retardation and ADHD. *Journal*


239. Swanson JM, Wigal SB, and Lemer MA. Comparison of the efficacy and safety of OROS methylphenidate HCl with methylphenidate tid and placebo in children with ADHD. *Pediatric Research.* 2000;47(4 suppl):34A.


Appendix E. Previous systematic reviews

Previous systematic reviews of this evidence are numerous.\textsuperscript{1-20} We included only four systematic reviews that we rated good quality\textsuperscript{14, 16, 20, 21}. The table below summarizes the characteristics and main findings of these four reviews. We rated the other reviews fair-poor quality primarily because they did not use standard methods of study appraisal. Also, many were not comprehensive in searching multiple databases and were nonspecific with regard to eligibility criteria and literature search strategies.

Inclusion criteria (study design, publication date, population characteristics, and interventions) and methods of analysis varied across the good-quality reviews. Despite this, main findings were generally consistent in suggesting that there are no clear differences in short-term efficacy and tolerability between MPH, DEX and pemoline. Additionally, the Jadad review (1999) summarized findings from longer-term, placebo-controlled trials of DEX and MPH that suggest these stimulants are associated with general improvement that persists over time.\textsuperscript{20} The Jadad review also summarized findings from placebo-controlled trials of MPH, antidepressants, pemoline, nicotine and phenylalanine in adults which suggested that the short-term efficacy of these treatments remained in question at that time.

Our review encompasses studies from all three good-quality reviews, as well as any published since 2001 and those that met our broader scope of interventions.

Summary of good quality systematic reviews

<table>
<thead>
<tr>
<th>Review</th>
<th>Characteristics</th>
<th>Main findings</th>
</tr>
</thead>
<tbody>
<tr>
<td>King 2004 (Centre for Reviews and Dissemination, Centre for Health Economics, University of York)</td>
<td>Study design: RCTs for efficacy/adverse events; systematic reviews for adverse events; Publication date: MPH=1999 and onward; DEX=1997 and onward; atomoxetine=1981 and onward; Population: Children and adolescents (≤ 18 years of age) diagnosed with ADHD (including hyperkinetic disorder); Interventions: MPH, DEX, atomoxetine; Total # of included studies: 65</td>
<td>In general, inadequate reporting of study methodology limited reliability of results. There was little evidence of consistent differences in short-term efficacy between MPH IR and ER, MPH IR and DEX IR, or MPH IR and atomoxetine. Adequate data regarding potential short-term adverse effects of MPH IR, MPH ER, DEX IR and atomoxetine is lacking.</td>
</tr>
<tr>
<td>Schachter 2001 (EPC at University of Ottowa)</td>
<td>Study design: Placebo-controlled RCTs; Publication date: 1981 or later; Population: ADD with or without hyperactivity; median age=8.7 years; Intervention: short-acting MPH; Total # of included trials: 62 (2897 patients)</td>
<td>Short-acting MPH demonstrated consistent short-term efficacy in reducing most ADD-related symptoms. Significant short-term harms reported by parents/patients included decreased appetite, insomnia, stomach ache, drowsiness and dizziness.</td>
</tr>
<tr>
<td>Jadad 1999 (EPC at McMaster University)</td>
<td>Study design: RCTs; Publication date: 1966 or later; Population: ADHD in humans; Interventions: DEX, MPH, pemoline, clonidine, bupropion,</td>
<td>Drug vs drug: There were few, if any differences in short-term efficacy between MPH, DEX and pemoline. Results of MPH and TCAs comparisons were conflicting. Body of drug vs drug evidence did not include any studies of clonidine, bupropion or SSRIs.</td>
</tr>
<tr>
<td>Review</td>
<td>Characteristics</td>
<td>Main findings</td>
</tr>
<tr>
<td>---</td>
<td>---</td>
<td>---</td>
</tr>
<tr>
<td></td>
<td>TCAs and SSRIs</td>
<td>Longer-term therapy (mean duration=20 weeks): Placebo-controlled trials of DEX or MPH in primarily school-age children suggest trends in general improvement over time regardless of treatment ADHD in adults: Short-term efficacy of MPH inconsistent across placebo-controlled trials Adverse effects: Short-term trials of stimulants most frequently examined sleep disorders/disturbances, headaches, motor tics, decreased appetite/anorexia, abdominal pain and irritability and no differences were reported. Nausea, fatigue and tiredness were also commonly examined and rates were similar for stimulants and antidepressants. Long-term safety data is inadequate to make any conclusions.</td>
</tr>
<tr>
<td></td>
<td>Total # of included trials (total # patients not reported): Drug vs drug=22 Long-term therapy=14 Treatment of ADHD in adults=12</td>
<td></td>
</tr>
<tr>
<td>Klassen 1998 Klassen 1999 (CCOHTA)</td>
<td>Study design: RCTs Publication date: 1981 or later Population: Children 0-18 years with diagnosis of ADD, ADDH or ADHD Intervention: DEX, MPH or pemoline for ≥ 1 week in duration Total # of included trials: 26 (999 patients)</td>
<td>No clear differences in short-term efficacy were found between MPH, DEX and pemoline. Safety: not reported</td>
</tr>
</tbody>
</table>

Reference