

How efficacious and safe is short-acting methylphenidate for the treatment of attention-deficit disorder in children and adolescents? A meta-analysis

Howard M. Schachter,* Ba' Pham,* Jim King,^{†‡}
Stephanie Langford,* David Moher*[‡]

Abstract

Background: Numerous small clinical trials have been carried out to study the behaviourally defined efficacy and safety of short-acting methylphenidate compared with placebo for attention-deficit disorder (ADD) in individuals aged 18 years and less. However, no meta-analyses that carefully examined these questions have been done. We reviewed the behavioural evidence from all the randomized controlled trials that compared methylphenidate and placebo, and completed a meta-analysis.

Methods: We searched several electronic sources for articles published between 1981 and 1999: MEDLINE, EMBASE, PsychINFO, ERIC, CINAHL, HEALTHSTAR, Biological Abstracts, Current Contents and Dissertation Abstracts. The Cochrane Library Trials Registry and Current Controlled Trials were also consulted. A study was considered eligible for inclusion if it entailed the following: a placebo-controlled randomized trial that involved short-acting methylphenidate and participants aged 18 years or less at the start of the trial who had received any primary diagnosis of ADD that was made in a systematic and reproducible way.

Results: We included 62 randomized trials that involved a total of 2897 participants with a primary diagnosis of ADD (e.g., with or without hyperactivity). The median age of trial participants was 8.7 years, and the median "percent male" composition of trials was 88.1%. Most studies used a crossover design. Using the scores from 2 separate indices, this collection of trials exhibited low quality. Interventions lasted, on average, 3 weeks, with no trial lasting longer than 28 weeks. Each primary outcome (hyperactivity index) demonstrated a significant effect of methylphenidate (effect size reported by teacher 0.78, 95% confidence interval [CI] 0.64–0.91; effect size reported by parent 0.54, 95% CI 0.40–0.67). However, these apparent beneficial effects are tempered by a strong indication of publication bias and the lack of robustness of the findings, especially those involving core ADD features. Methylphenidate also has an adverse event profile that requires consideration. For example, clinicians only need to treat 4 children to identify an episode of decreased appetite.

Interpretation: Short-acting methylphenidate has a statistically significant clinical effect in the short-term treatment of individuals with a diagnosis of ADD aged 18 years and less. However, the extension of this placebo-controlled effect beyond 4 weeks of treatment has not been demonstrated. Exact knowledge of the extent and definition of the short-term behavioural usefulness of methylphenidate is questioned.

Studies across North America have shown that attention-deficit hyperactivity disorder (ADHD) affects 3%–5% of children aged 18 years and less, making it perhaps the most common psychiatric diagnosis in this age group.^{1–17} Short-acting methylphenidate (Ritalin) is the medication that is almost universally pre-

Research

Recherche

From *the Thomas C. Chalmers Centre for Systematic Reviews, Children's Hospital of Eastern Ontario (CHEO) Research Institute; †the Department of Pediatrics, CHEO; and ‡the Department of Pediatrics, University of Ottawa, Ottawa, Ont.

This article has been peer reviewed.

CMAJ 2001;165(11):1475-88

scribed for ADHD in these children,^{4,10,18–22} making it the de facto “gold standard.”^{7,5,10,15,23,24}

A large number of relatively small randomized controlled trials (RCTs) have examined the effect of this central nervous system stimulant on the core behavioural features of ADHD, namely, age-inappropriate levels of inattention, impulsivity and hyperactivity.^{1,5,20–22,25–27} Several meta-analyses have synthesized this behavioural evidence,^{2,3,28–41} yet each of these is flawed.⁴² For example, they did not investigate adequately safety data, the impact of sources of clinical heterogeneity or the presence of publication bias.⁴¹ Few satisfactorily distinguished among the various types of stimulant used,^{38–41} despite evidence for their different pharmacokinetic profiles, clinical regimens, responses and risks (e.g., the liver toxicity of pemoline).⁴³

More important, most focused on the question of efficacy of stimulants relative to other treatments (e.g., behavioural therapy).^{39,41} Few looked exclusively at the clinical utility of methylphenidate compared with placebo.^{3,42} This is noteworthy, because comparing a drug with placebo is essential to understanding whether or not it works and is safe.⁴⁴ A given intervention may work better than another one, without either of them being significantly better than no active intervention at all. Results from placebo-controlled studies provide a meaningful context in which to interpret evidence concerning a drug’s efficacy relative to that of other approaches to clinical care.

We performed a meta-analysis that took into account possible population, intervention and outcome sources of heterogeneity, including differing primary diagnoses, sex, cognitive-developmental level or age, dose, treatment duration and the use of co-interventions. In addition, we investigated the robustness and validity of the effect of methylphenidate in light of trial quality, study design and publication bias. All analyses were planned. As ADHD is not a single diagnostic entity,^{5,8,10,21,41,45–47} the term “attention-deficit disorder” (ADD) is employed to refer to the entire range of possible forms of the disorder (e.g., with or without hyperactivity).

Methods

Without restriction on either the publication or language status of reports, we searched several electronic sources: MEDLINE (1981–December 1999), EMBASE (1988–November 1999), PsycINFO (1981–November 1999), ERIC (1981–September 1999), CINAHL (1982–September 1999), HEALTHSTAR (1981–November 1999), Biological Abstracts (1998–September 1999), Current Contents (1997–November 1999) and Dissertation Abstracts (1990–October 1999). Our searches included a standardized filter to capture RCTs and focused on the following terms in the titles, abstracts, and key word lists of all citations: “attention deficit disorder,” “attention deficit disorder with hyperactivity,” “hyperactivity,” “hyperkinesis,” “hyperkinesia,” “methylphenidate,” “Ritalin” and “psychostimulants” (Appendix 1).

The Cochrane Library’s Trials Register (1981–December 1999) and Current Controlled Trials (January–August 1999) were

also consulted. The former constituted a surrogate manual search. Reference lists from the RCTs that were included in the meta-analysis and from pertinent reviews were searched manually, as were the files of content experts.

Trial selection

We considered a trial eligible for inclusion if it was a placebo-controlled RCT involving short-acting methylphenidate given to children aged 18 years or less who had received a primary diagnosis of ADD made in a systematic and reproducible way. Grounds for exclusion included the following: reports published earlier than 1981, that is, before the likely influence of the newly published *Diagnostic and Statistical Manual of Mental Disorders (DSM-III)*⁴⁸ criteria that were better differentiated and more stringent than those used previously to identify ADD in its various forms (e.g., with or without hyperactivity); *n*-of-1 studies; trials with participants who had experienced medical or psychiatric conditions that required a highly specialized school or home environment, or both (e.g., mental retardation, autism, psychosis); and studies with participants who were receiving stimulants other than methylphenidate.

Two assessors independently screened the title, abstract and key words for each citation to determine whether to retain it. Potentially relevant citations were retrieved and then subjected to a relevance assessment using our inclusion and exclusion criteria. A reliability study with 20 randomly selected, potentially relevant full-text articles achieved 75% agreement. The remaining full-text documents were then independently assessed, and the reasons for excluding reports were noted. Disagreements were settled by forced consensus.

Trial quality assessment

Using 2 measures, one of us (D.M.) assessed for quality all reports that referred to a trial that was included in the meta-analysis. One was the validated, 3-item Jadad scale for randomization (0–2 points), double-blinding (0–2) and the description of withdrawals and dropouts (0–1);⁴⁹ the sum of these scores yields a Jadad total score in which 3 and above is considered high quality. The other was an index of the concealment of treatment allocation (i.e., adequate, inadequate or unclear).⁵⁰

Data abstraction

Two of us (B.P. and S.L.) abstracted the data using a structured form created a priori to capture report (e.g., language of publication), trial (e.g., design), population (e.g., primary diagnosis) and intervention (e.g. dose) characteristics, as well as behavioural efficacy and adverse event data (e.g., insomnia.) Reports were not masked.⁵¹

It was virtually impossible to extract an estimate of the treatment effect from a proper analysis of variance for crossover trial data, because of the different ways these trials reported their data. However, the authors did consistently report the means and standard deviations for the efficacy measures pertaining to both methylphenidate and placebo. These were the summary data that we extracted and used in efficacy analyses. Assumptions that were verified later in sensitivity analyses were made regarding the absence of a crossover effect and the independence of cohorts.

We employed 2 strategies exclusively to deal with primary outcomes. For trials that failed to report variance measures of effect, these data were imputed from the reported *p* values of the *t*-test of treatment comparisons or the *F* test of the treatment effect from the pertinent analysis of variance.⁵² Trial results that were exclusively expressed graphically were scanned, the coordinates of various data points were extracted, and these data were summarized.⁵³

Derivation of summary measures

To deal with the possibility of different, albeit similarly scaled (e.g., continuous data only), instruments referring to a single feature of ADD, or with slightly different versions of a given scale (e.g., the hyperactivity index²⁵), we derived the effect size (i.e., standardized mean difference) for all efficacy outcomes. For trials with multiple doses or stratification, an average effect size was computed across the dose levels or strata. For example, a crossover trial with twice-a-day dosing might have provided summary data on efficacy outcomes with participants receiving 0.15 mg/kg, 0.50 mg/kg or 0.75 mg/kg of methylphenidate. A random effects model was employed to combine the summary data across the 3 dose levels into an average effect size, taking into account both the within-dose and between-dose variations. The variance of the average effect size was considered "within" variation (i.e., variance of a treatment effect estimate from a fixed effects model). The average effect sizes for these trials were used in evaluating the global picture of methylphenidate. Where intervention length was expressed as a range (e.g., 7–10 days), the lower bound was entered into analyses.

Differences in responder numbers between treatment phases were most likely to be observed in trials with a crossover design. Thus, the proportions of clinical response (e.g., a difference in treatment-phase responders expressed as a proportion of the number of patients in a crossover trial) were combined across trials. The proportions of patients who experienced side effects (decreased appetite, insomnia, headache, stomach ache, drowsiness, anxiety and dizziness) that were highlighted often in clinical and empirical work^{18,43,54} were captured for the treatment phases or conditions.

Outcomes

Ratings by parents and teachers of efficacy were used, because they were well situated to evaluate overt behaviour. Given their popularity in methylphenidate trials for ADD, the teacher and parent versions of the hyperactivity index (HI) became the primary outcomes (HI-T, HI-P) by default. Parent and teacher rating data were also abstracted for other global, core and related features of ADD.

We included self-reported indices of side effects to compensate for the possible underestimation of these events by external observers.¹⁸ Different definitions provided by trialists of clinical response to methylphenidate were used in the trials included in the meta-analysis. We did not attempt to reconcile the various definitions. Given the scarcity of safety data, we used trialists' definitions of "side effect" and "serious side effect" in the present analyses without attempting to reconcile either set of definitions.

A given trial could contribute reported data from any number of teacher-reported or parent-reported behavioural instruments respectively. Our goal was to combine data derived from the same instrument, if possible. Yet, in certain cases, and across studies, data from different instruments (e.g., 2 versions of the Conners

scales) that indexed the same construct (e.g., hyperactivity) were synthesized. No single trial that reported data from 2 or more instruments that measured the same construct had all of these data points entered into the meta-analysis. Our choice of which data to synthesize was made based on the best congruence regarding construct, and in terms of scaling (e.g., continuous data only), with the instruments that had provided data in the other studies. Space limitations preclude our definition of the instruments.

Analysis of efficacy data

All of our analyses directly compared methylphenidate with placebo according to the intention-to-treat principle and included data that reflected the last follow-up for participants who were receiving treatment. Trial effect sizes or average effect sizes for stratified trials were combined using a random effects model. We also used this same approach to combine the treatment differences with respect to the proportions of clinical responders.

Analysis of adverse event data

We employed an approach to multitreatment trials similar to that of Hasselblad⁵⁵ whereby the percentage of patients experiencing a side effect or serious side effect was treated as a dependent variable in a random effects analysis of variance. For each trial, a covariance matrix was constructed assuming a correlation of 0.5 for the percentages of patients with side effects while they were on methylphenidate and placebo within a crossover design. Treatment and trial factors were the independent variables in the random effects model.

We fitted the model using an approach similar to that of Normand.⁵⁶ Heterogeneity between trials was assessed by including a treatment-by-trial interaction. The proportions of patients who reported a side effect by treatment condition, in addition to treatment differences, were derived from the fitted model, along with 95% confidence intervals. We then used the estimated proportions to derive a relative risk of the side effect. The associated number needed to harm was derived using the estimated relative risk and interpreted against the median control group rate. The number needed to harm refers to the number of individuals receiving treatment before one of them is observed to be experiencing an adverse event.

Sensitivity analyses

We planned all sensitivity and subgroup analyses with primary outcomes, yet the exact data organization scheme employed in each was data driven. Our sensitivity analyses investigated 2 facets of trial quality (Jadad total score, allocation concealment) as well as study design (i.e., crossover v. parallel trials). To compensate for the correlation in measurements from different crossover phases of a trial potentially leading to an underestimate of variance, each trial's variance was reduced by 20% and the treatment effects incorporating the adjusted variances were re-estimated using a random effects model.

Subgroup analyses

Our subgroup analyses assessed the impact of intervention length (≤ 1 week v. 2–4 weeks v. > 4 weeks), co-interventions

and dose. As there was considerable variability in the dose of methylphenidate, including within a given day, we employed dose “potency” (e.g., 0.6 mg/kg twice a day = 0.6 mg/kg 3 times a day) rather than the total or mean daily dose. Where expressed in absolute terms, we translated dose into mg-per-kg values for a sample by way of converting the sample’s mean age to a mean weight.⁵⁷ This afforded comparability across studies.⁴¹ The following scheme for dose potencies was adopted: low (≤ 0.3 mg/kg), medium ($0.3 \text{ mg/kg} < x < 0.75 \text{ mg/kg}$) and high ($\geq 0.75 \text{ mg/kg}$).

Analysis by cognitive-developmental stage was prevented by the state of the reported data in the RCTs. Instead, we developed an ad hoc scheme based on age that ignored many key developmental epochs (i.e., ≤ 12 v. > 12 years of age). Because few studies differentiated data by sex, we decided to compare efficacy estimates from those trials that had a homogeneous distribution of males with those that included some female participants. Finally, given that diagnostic systems use different criteria for a given diagnosis, we chose to perform separate analyses of trial data differentiated by primary diagnosis (and system): for example, ADD with hyperactivity (ADHD, i.e., DSM-III), ADHD (i.e., DSM-III-Revised)⁵⁸ and an undefined mix of ADD diagnoses from any system. Given the status of the trial data, we were unable to undertake separate analyses for different combinations of primary diagnosis and comorbidity.

Publication bias

Our assessment of publication bias⁵⁹ entailed the visual inspection of funnel plots of effect sizes versus their precision. Egger and colleagues⁶⁰ graphical test was used to evaluate the degree of asymmetry in the funnel plots. Duval and colleagues⁶¹ “trim and fill” method was used to estimate the number of unobserved trials and derive the treatment effect estimates adjusted for publication bias.

Results

Of 487 citations entered into broad screening, 224 were considered potentially relevant; their full reports were re-

trieved, and these were assessed formally for relevance. Reasons for excluding reports⁶² ($n = 145$) from this last phase included nonbehavioural outcome (75), no randomization (27), an unsystematic (e.g., a single scale) diagnostic procedure (17), no placebo control (9), participants received an excluding co-diagnosis (7), some participants received stimulants other than methylphenidate (5), n -of-1 study (2), follow-up report (2) and not all participants received an ADD diagnosis (1).

From the remaining 79 relevant reports, we identified 62 unique trials. All reports were published between January 1981 and December of 1999 and exclusively in English-language journals. These reports are listed in Appendix 2 (available on the *CMAJ* Web site at www.cma.ca/cmaj/vol-165/issue-11/pdf/deficitappendix.pdf). Crossover designs predominated (83.9%), yet only 13.5% of these evaluated a carryover effect.

The trials that we reviewed included 2897 participants (Tables 1–6). Only 11% of studies randomized more than 80 participants; the mean sample size was 46.7 (range 11–234). The median age of all trial participants was 8.7 years (range 2.4–18 years); these data are taken from 52 trials. The median “percent male” composition for the trials included in the meta-analysis was 88.1% (range 0%–100%); these data are taken from 59 trials.

The number of trials that included participants with a homogeneous primary diagnosis was 45 (72.6%); they included the following: ADHD (38.7%) and ADHD

Table 2: % of diagnostic systems employed*

- DSM-III†: 45.1 ($n = 28$)
- DSM-III-R‡: 38.7 ($n = 24$)
- Test battery with predefined research diagnostic criteria: 9.7 ($n = 6$)
- DSM-IV§: 6.5 ($n = 4$)

*No. of trials = 62.
†Diagnostic and statistical manual of mental disorders (DSM-III).⁴⁸
‡Diagnostic and statistical manual of mental disorders (DSM-III-R).⁵⁸
§Diagnostic and statistical manual of mental disorders (DSM-IV).¹

Table 1: % of trials ($n = 62$) organized by primary diagnoses received by participants

- ADHD: 38.7 ($n = 24$)
- ADHD: 33.9 ($n = 21$)
- ADHD + ANX v. ADHD – ANX: 3.2 ($n = 2$)
- ADD + aggression/ oppositionality v. ADD – aggression/ oppositionality v. ADD undefined: 1.6 ($n = 1$)
- ADD mixed (ADHD, ADD-H): 9.7 ($n = 6$)
- ADD mixed (combined v. inattentive type): 3.2 ($n = 2$)
- ADHD undefined: 3.2 ($n = 2$)
- ADD undefined (likely ADHD): 1.6 ($n = 1$)
- ADD (likely ADHD): 1.6 ($n = 1$)
- “Hyperactive” via ADD core features (likely ADHD): 1.6 ($n = 1$)
- ADD undefined: 1.6 ($n = 1$)

Note: ADHD = attention-deficit hyperactivity disorder, ADHD = attention-deficit disorder, with hyperactivity, ANX = anxiety disorder, ADD = attention-deficit disorder, ADD-H = attention-deficit disorder, without hyperactivity.

Table 3: % of trials ($n = 37$) in which a comorbid condition was identified, organized by the 2 most frequent comorbidities per trial

- EXT: 40.6 ($n = 15$)
- EXT, INT: 18.9 ($n = 7$)
- EXT, LD: 5.5 ($n = 2$)
- EXT, DD: 2.7 ($n = 1$)
- LD, EXT: 5.5 ($n = 2$)
- TIC, EXT: 5.5 ($n = 2$)
- INT, EXT: 8.1 ($n = 3$)
- TIC: 2.7 ($n = 1$)
- LD: 5.5 ($n = 2$)
- INT: 5.5 ($n = 2$)

Note: EXT = externalizing disorder, INT = internalizing disorder, LD = learning disorder, DD = developmental disorder (other than excluding diagnoses), TIC = major tic disorder (e.g., Tourette’s syndrome).

(33.9%). The DSM-III-R and DSM-III diagnostic systems guided 38.7% and 45.1% of trial methodologies respectively.

In 59.7% of the trials, a comorbid condition was identified. Of these, externalizing disorders were either the most (67.7%) or second most frequently diagnosed comorbidity (19.1%). In 19 trials, patients received a homogeneous primary diagnosis without any comorbidity being reported or evaluated (i.e., 12 and 7 with ADDH and ADHD respectively).

Interventions lasted, on average, 3.3 weeks (range 2 days to 28 weeks). Only 9 RCTs (14.5%) had interventions that lasted longer than 4 weeks, and more than half of the 62 trials lasted no more than 10 days. One study lasted 12 months, yet only a subset of participants was followed beyond 4 months.^{63,64}

The most frequently used dose-related conventions included the following: 5-mg, 10-mg, 15-mg versus 20-mg dose potency contrast ($n = 6$ trials), 0.3 mg/kg dose potency (irrespective of potency contrasts) ($n = 26$), and a twice-a-day dosing schedule ($n = 39$). Across the 62 RCTs, 38 different types of methylphenidate intervention (i.e., combinations of dose potency contrast by schedule: for example, 0.3 mg/kg twice a day versus 0.6 mg/kg twice a day) were employed.

A co-intervention (e.g., a summer treatment program, with behavioural regimens) was provided in 25.8% of the studies, whereas in 11.3% of the RCTs, participants continued to receive a pretrial intervention (e.g., a specific special education service).

Trial quality

No trial report received a maximum 5-point Jadad total quality score, 5 (8.1%) studies achieved a score of 4, 38.7% were judged to have low quality (0–2 points), and the mean total quality score was 2.6 (low). Only 5 (8.1%) trials reported adequate allocation concealment, with the remaining descriptions deemed “unclear.”

Table 4: Most frequent “clinical pictures” (i.e., combinations of primary diagnosis and comorbidity)*

- ADDH alone ($n = 12$)
- ADHD alone ($n = 7$)
- ADD or hyperactive (likely ADDH) alone ($n = 2$)
- ADHD + ODD or (ODD or CD) ($n = 11$), CD ($n = 2$), TIC ($n = 3$) or ANX ($n = 1$)
- ADDH + ODD ($n = 5$), LD ($n = 3$) or aggression ($n = 1$)
- ADHD +/- ANX ($n = 2$)
- ADD +/- aggression/oppositionality ($n = 1$)
- ADD mixed + no comorbidity ($n = 2$), ODD ($n = 2$), ODD or CD ($n = 1$), INT ($n = 2$) or LD ($n = 1$)
- ADD or ADHD undefined + no comorbidity ($n = 2$) or ODD ($n = 2$)

Note: ODD = oppositional–defiant disorder, CD = conduct disorder.
*No. of trials = 62.

Analyses of the primary efficacy outcomes

HI-T data (effect size 0.78, 95% CI 0.64–0.91) and HI-P data (effect size 0.54, 95% CI 0.40–0.67) exhibited significant effects in favour of methylphenidate (Figs. 1 and 2). This represents a decrease of 6 and 4 points on the respective HI scales from a mean of 14 in patients who received placebo within the largest trial of 161 participants.⁶⁵ These changes are “interpreted” in terms of the largest trial that provided HI data, because it is reasonable to assume that it would be the least biased.

Analyses of additional efficacy outcomes: core features, clinical response and related features data

Teacher-reported results similar to those seen for the primary outcome were recorded for clinical response (teacher or staff), global indices, core features and key externalizing features (Fig. 1). However, the magnitudes of effect were variable and tended to be smaller than for the primary outcome; for both attention and emotional lability, the effect was not statistically significant.

Methylphenidate exhibited significant, though variable and weaker, effects on parent-reported clinical response, global indices, core features and key externalizing features. Estimates regarding inattention, hyperactivity/impulsivity and oppositional–defiant behaviour were not statistically significant.

Minimal evidence of statistical heterogeneity was observed. Only 4 of 16 teacher outcome analyses (Fig. 1) and 3 of 13 parent outcome analyses (Fig. 2) respectively were associated with significant levels of statistical heterogeneity. Insufficient comparability in tic-related outcomes precluded the analysis of teacher and parent data.

Sensitivity analyses

We are contrasting the results reported by teachers and those reported by parents (Figs. 3 and 4). Even though the

Table 5: Most frequently used dose-related conventions: (A) dose potency contrasts, (B) single potency, (C) dose schedule*†

- (A) 5, 10, 15, 20 mg ($n = 6$), 0.3, 0.6 mg/kg ($n = 5$), 0.3 mg/kg ($n = 5$), 0.3, 0.5 mg/kg ($n = 4$)
- (B) 0.3 mg/kg ($n = 26$), 10 mg ($n = 17$), 0.5 mg/kg ($n = 11$)
- (C) Twice a day ($n = 39$), once a day ($n = 13$)

*38 different combinations of dose potency contrast and schedule.
†No. of trials = 62.

Table 6: % of trials with a co-intervention and % of trials with participants continuing with a pre-existing intervention

- 25.8 (16/62): with a summer treatment program (50%, 8/16)
- 11.3 (7/62): with various special education services (57.1%, 4/7) or nonstimulant drug therapies (42.9%, 3/7)

changes in methylphenidate effect were more pronounced for teacher-reported data, trial quality and efficacy were inversely related for both informants. The initial HI-T and HI-P estimates of effect decreased by 9.0% and 5.6% for high-quality trials respectively, yet increased by 26.9% and 24.1% for low-quality studies. Reports with adequate allocation concealment increased the methylphenidate effect by 39.7% and 13% for teacher and parent data respectively. All effects remained statistically significant.

Our analyses of study design revealed that only the 2 methylphenidate effects for crossover trials were statistically significant (Figs. 3 and 4). When we reanalyzed both teacher and parent data following the variance adjustment for crossover designs, the estimates remained virtually unchanged. Only the adjusted HI-T estimate displayed evidence for statistical heterogeneity.

Subgroup analyses

Changes tended to be more pronounced for teacher data, and most remained statistically significant (Figs. 3 and 4). The original HI-T estimates increased for both age categories, and the HI-P estimates increased only for older participants. For both informant sources, trials exclusively including males exhibited a stronger methylphenidate effect than those in which both sexes were represented. Only the HI-P effect for mixed-sex trials did not increase its initial methylphenidate estimate.

For all categories of primary diagnosis, and for both informants, we observed statistically significant methylphenidate effects. Both the HI-T and HI-P estimates for ADHD diagnosis studies decreased their initial methylphenidate effect estimates. The strongest teacher

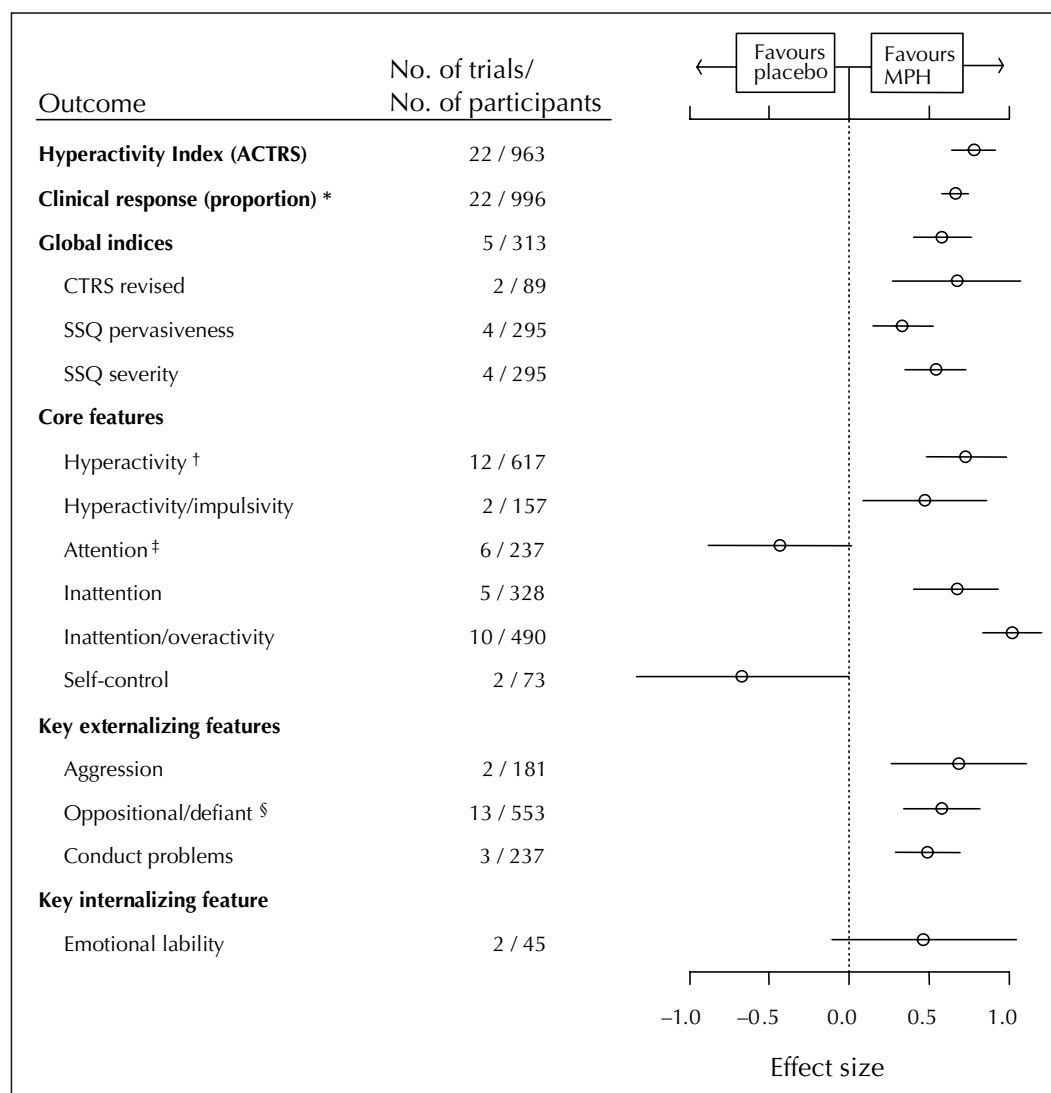


Fig. 1: Effect sizes for the hyperactivity index: teacher ratings (95% confidence interval [CI]). MPH = methylphenidate, ACTRS = Abbreviated Conners Teacher Rating Scale, CTRS = Conners Teacher Rating Scale, SSQ = School Situations Questionnaire. *p* values for statistical heterogeneity: **p* < 0.001, †*p* = 0.03, ‡*p* = 0.02, §*p* = 0.03.

and parent effects were seen for ADHD and mixed diagnosis trials respectively.

For teacher data, methylphenidate effect magnitude and dose potency appeared to be positively related (Fig. 3), suggesting a possible trend.

Each of the HI-T and HI-P methylphenidate effect sizes was positively related to intervention length for interventions lasting no more than 4 weeks. In both cases, the 1-week and 2–4-week duration were decreased and increased from their initial estimates respectively. The trials that lasted more than 4 weeks entailed parallel designs.

The teacher-reported methylphenidate effects for trials with and without a co-intervention decreased and slightly increased the original methylphenidate effect respectively. Trials without a co-intervention maintained the original HI-P estimate.

Two instances of HI-T-related statistical heterogeneity were observed in analyses involving dose potency. No HI-P subgroup analyses were associated with evidence for statistical heterogeneity (Figs. 3 and 4).

Publication bias

The rank correlation test revealed significant inverse relations for HI-T ($r = -0.45$, $p = 0.004$) and HI-P ($r = -0.36$, $p = 0.04$). The graphical test confirmed substantial funnel plot asymmetry (Fig. 5) for HI-T (asymmetry 1.07, $p < 0.001$) and HI-P (asymmetry 0.72, $p < 0.001$). Regarding teacher data, it was estimated that 8 trials (95% CI lower bound 5) showing a lesser or nonsignificant methylphenidate effect might have been suppressed (Fig. 5A). The adjustment decreased the initial teacher estimate by 21% (to 0.62,

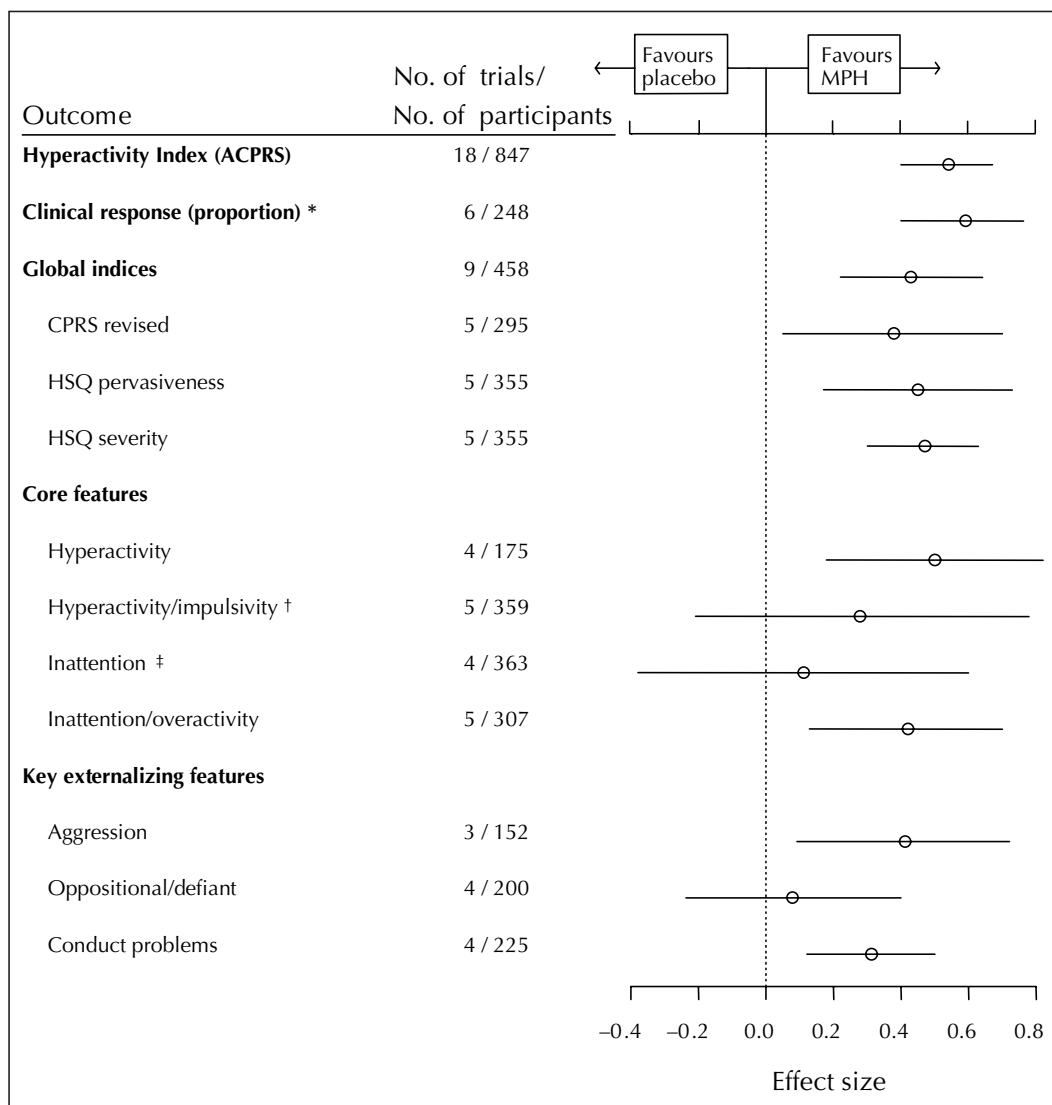


Fig. 2: Effect sizes for the hyperactivity index: parent ratings (95% CI). ACPRS = Abbreviated Conners Parent Rating Scale, CPRS = Conners Parent Rating Scale, HSQ = Home Situations Questionnaire. p values for statistical heterogeneity: * $p < 0.001$, † $p = 0.001$, ‡ $p = 0.008$.

Fig. 3). For HI-P data, evidence of publication bias lead to a 3.7% decrease in the original estimate (to 0.52, Fig. 4). It was estimated that one trial (95% CI lower bound 0) that showed a lesser or nonsignificant methylphenidate effect might have been suppressed (Fig. 5B).

Analyses of adverse events

We observed that almost all the differences between methylphenidate and placebo were in favour of a higher percentage of adverse events while participants were taking methylphenidate (Table 7). The largest statistically significant divergences were exclusively in favour of methylphenidate and entailed parent/self ratings: all related

decreased appetite (30.3%, 95% CI 18.0–42.6), insomnia (17.0%, 95% CI 8.3–25.8) or stomach ache events (9.0%, 95% CI 1.2–16.9) and serious decreased appetite events (8.7%, 95% CI 3.6–13.9). Teacher/staff ratings for serious decreased appetite events (6.1%, 95% CI 0.2–12.0) also showed a statistically significant difference in favour of methylphenidate. Children were more likely to experience anxiety (reported by parent/self) and headache (reported by teacher/staff) events while on placebo,⁵⁴ although neither of these results achieved statistical significance.

Overall, the number-needed-to-harm data for “all related” adverse events exhibited a wide variation (Table 7): for example, 4 (decreased appetite) to 22 (headache) for parent/self data compared with 40 (decreased appetite) to

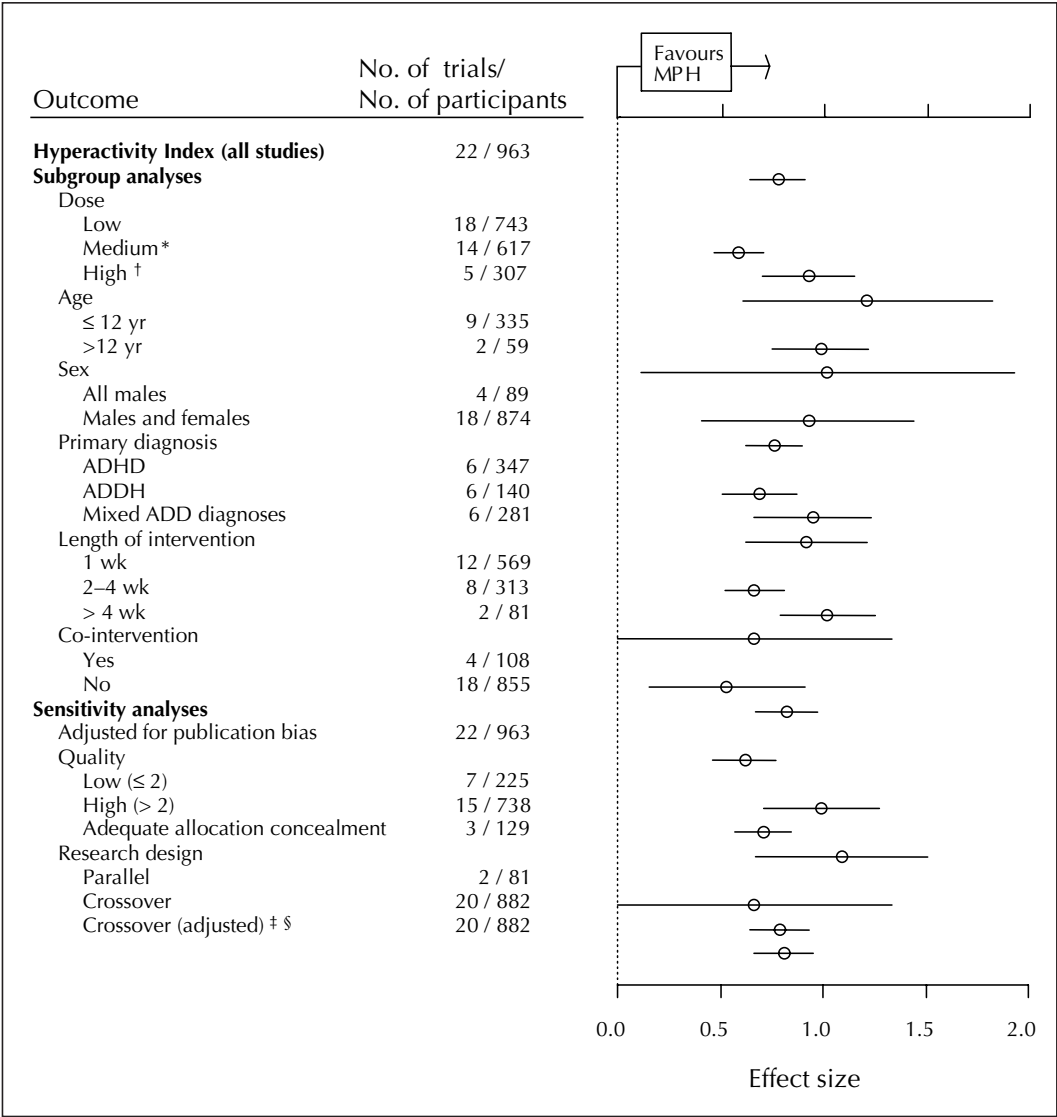


Fig. 3: Sensitivity and subgroup analyses of the hyperactivity index: teacher ratings (95% CI). ADHD = attention-deficit hyperactivity disorder (DSM-III-Revised),⁵⁸ ADDH = attention-deficit disorder with hyperactivity (DSM-III).⁴⁸ *p* values for statistical heterogeneity: **p* = 0.006, †*p* < 0.001, ‡*p* = 0.04. §Variance from each trial reduced by 20% to compensate for correlation in crossover phases potentially leading to an underestimate of variance.

367 (drowsiness) for teacher/staff data. For serious adverse events, a narrower range characterized the parent/self-related number-needed-to-harm estimates (i.e., 14 [decreased appetite] to 26 [insomnia]); and, only one teacher/staff-indexed number needed to harm (i.e., decreased appetite) could be derived.

Prominent number-needed-to-harm results were derived exclusively from parent/self ratings of 5 specific adverse events (Table 7). Only 4 and 7 study participants receiving methylphenidate, respectively, were required for a decreased appetite and insomnia-related event to be identified. Three other important number-needed-to-harm values were observed: all stomach ache events (number needed to harm = 9), all drowsiness events (number needed to harm = 10) and all dizziness events (number needed to harm = 11).

We found no evidence of statistical heterogeneity asso-

ciated with any of the methylphenidate effects or prominent number-needed-to-harm values.

Interpretation

We have shown that short-acting methylphenidate quickly and efficaciously reduces most of the clinical manifestations of ADD in children aged 18 years and less. Teacher and parent hyperactivity index scores were reduced by 6 (i.e., 42.9%) and 4 points (i.e., 28.6%) respectively. However, we cannot interpret these primary outcome results meaningfully, because the normative data for this measure are age-dependent and sex-dependent,²⁵ whereas the study used to "interpret" them included an undifferentiated mix of participants varying on these bases.⁶⁵

Although these results are impressive, we are uncertain

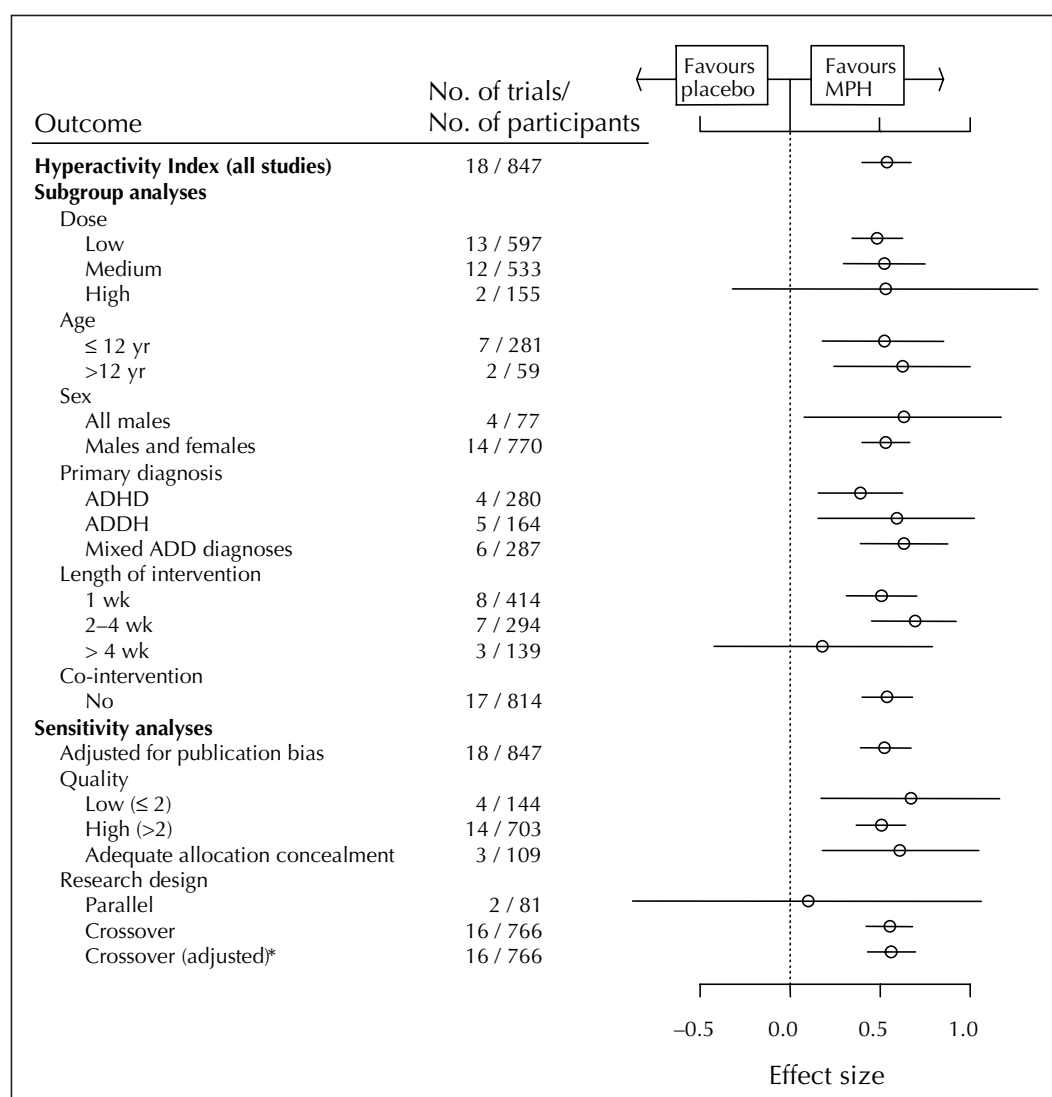


Fig. 4: Sensitivity and subgroup analyses of the hyperactivity index: parent ratings (95% CI). *Variance from each trial reduced by 20% to compensate for correlation in crossover phases potentially leading to an underestimate of variance.

about whether they are robust or completely valid. For example, collectively these trials were characterized by low quality defined 2 ways; and low Jadad-defined quality inflated our hyperactivity index estimates.⁵⁰ We also detected a substantial amount of publication bias that, when used to adjust the estimates of efficacy, decreased the teacher-defined hyperactivity index estimate by 21%. Other statistically significant estimates for teacher-rated core features might similarly decrease if investigated likewise. The indication of publication bias suggests that studies showing no effect, or involving children on methylphenidate who fared less well than those receiving placebo, may not have been published. We emphasize the teacher data, because these informants were probably best placed to observe the impact of the most frequently employed dosing regimens (i.e., 52 of 62 trials used once-a-day or twice-a-day dosing).

We also observed that the core feature and global index

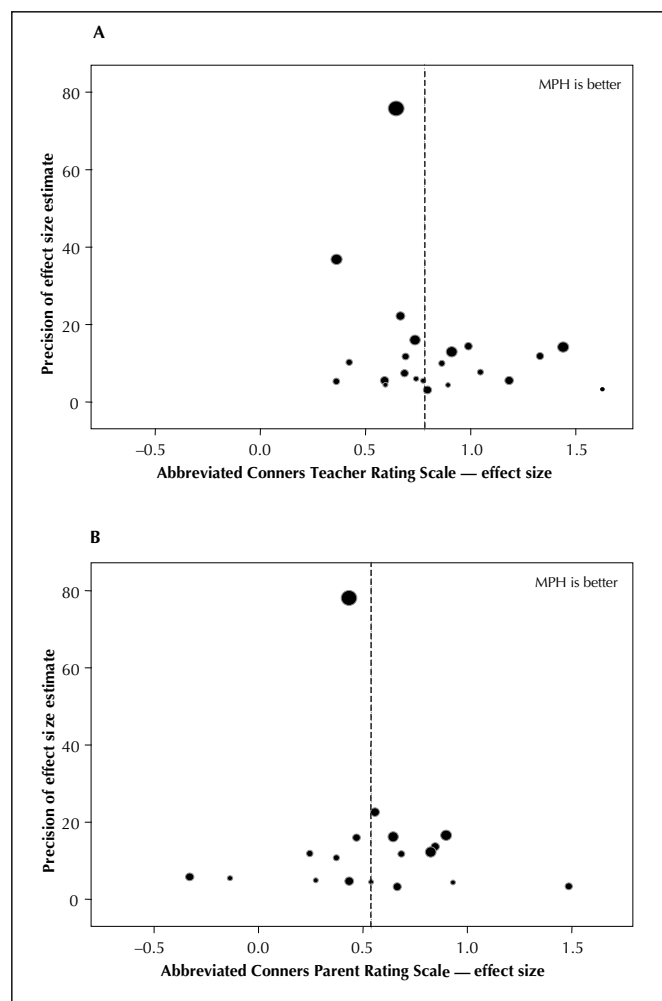


Fig. 5: (A) Funnel plot of methylphenidate effect size for the hyperactivity index (T) versus its precision. The trials' symbols are proportional to sample size (median 32, range 11–161). (B) Funnel plot of methylphenidate effect size for the hyperactivity index (P) versus its precision. The trials' symbols are proportional to sample size (median 37, range 11–161).

results rarely exhibited the same magnitude as the hyperactivity index–defined effects, if any effect at all (e.g., teacher-related attention, parent-rated inattention and hyperactivity/impulsivity). This supports the view of the inconsistency in definition, or “outcome dependence,” of the effect of short-term methylphenidate.^{18,22,66}

Moreover, we were unable to demonstrate that the methylphenidate effect is maintained beyond 4 weeks. Few trials (14.5%) included treatments of this duration; perhaps due to the dearth of trials and participants, both hyperactivity index–defined methylphenidate effects for these longer-term, exclusively parallel design trials (maximum 28 weeks) failed to achieve statistical significance. The paucity of long-term trials^{3,10,20–22,40} is problematic, because children routinely receive methylphenidate in clinical contexts for much longer than was observed in the present collection of small trials.^{7,67} The recently completed MTA trial,⁶⁸ though lacking a placebo group, may address some of the concerns identified in this meta-analysis regarding long-term treatment. It was methodologically sound and had sufficient power to detect the superiority of medical management (with 73.4% of participants maintained on methylphenidate at study end) over a relatively long-term period (i.e., 14 months) in a comprehensively assessed patient population.

We also found that whereas adverse event data were underreported by trialists,^{69–71} both parent/self and teacher/staff data revealed that serious episodes of decreased appetite were statistically significantly more common when children were on methylphenidate rather than placebo (Table 7). Significant differences in favour of methylphenidate also characterized other parent/self-observed adverse events (i.e., all related instances of decreased appetite, insomnia, stomach ache, headache, dizziness, and serious headache and stomach ache events). Although requiring cautious interpretation,⁷² notable parent/self-derived number-needed-to-harm data highlighted difficulties with decreased appetite and insomnia and, to a lesser extent, stomach ache, drowsiness and dizziness. Although methylphenidate-related adverse events can be dose-dependent and can diminish over time,^{18,21,22} the dearth of data precluded the evaluation of the impact of either of these factors on the safety profile. Finally, as with efficacy data, we do not know whether this short-term safety profile²¹ persists in the long term. The longer-term MTA trial results highlighted similar safety issues, in that 64.1% and 14.3% of children who received a stimulant exhibited side effects of any severity or of a moderate-to-severe kind respectively.⁶⁸

Overall, the present results confirm the findings from other meta-analyses of a short-term “methylphenidate effect.”^{23,41,42} However, in light of the observed publication bias, the influence of trial quality on efficacy estimates and the outcome dependence of the methylphenidate effect, our results cannot be used to confirm suggestions in previous meta-analyses that we know the *exact extent to which*, or *in exactly which terms*, short-acting methylphenidate is effi-

cacious in the short-term. This second finding may or may not be surprising given that, unlike other attempts, our question was very specific (i.e., methylphenidate v. placebo), our inclusion criteria did not limit the number or types of relevant behavioural outcome and, in evaluating efficacy data, we explored issues of clinical heterogeneity, robustness and validity (e.g., publication bias, trial quality).^{3,42} Our safety-related findings cannot be compared with those of other meta-analyses, because the latter invariably did not investigate these data systematically.^{3,39,41}

Limitations of our work include debatable decisions to use age as a surrogate for cognitive-developmental level and to derive a dose potency typology. Certain results, however meaningful (e.g., positive relation between dose magnitude and efficacy), are likely to be uninterpretable. In addition, the innovations that dealt with crossover trials probably require further validation despite the negligible change in estimates yielded by reanalysis. Although instances of statistical heterogeneity were unexpectedly rare,³ this finding may have resulted from underpowered statisti-

Table 7: Adverse events experienced by participants receiving methylphenidate or placebo

Adverse event	Assessors	No. of trials/ no. of pts	MPH, % (95% CI)	Placebo, % (95% CI)	Difference (95% CI)	NNH (median placebo event rate, %)	<i>p</i> value*
Decreased appetite							
All related events	Parent/ Self	10/675	44.8 (36.8–52.7)	14.4 (5.1–23.8)	30.3 (18.0–42.6)	4 (11.3)	0.11
	Teacher/ staff	3/129	46.0 (38.7–53.3)	43.6 (35.3–51.9)	2.4 (–5.5 to 10.3)	40 (45.5)	0.34
Serious side effects	Parent/ self	5/222	10.5 (7.0–14.0)	1.8 (0.0–5.6)	8.7 (3.6–13.9)	14 (1.5)	0.27
	Teacher/ staff	3/129	26.4 (21.5–31.3)	20.3 (14.6–26.1)	6.1 (0.2–12.0)	175 (1.9)	0.66
Insomnia							
All related events	Parent/ self	9/663	47.7 (42.1–53.3)	30.7 (23.9–37.5)	17.0 (8.3–25.8)	7 (27.1)	0.10
Serious side effects	Parent/ self	5/222	10.9 (7.5–14.3)	7.1 (3.0–11.3)	3.8 (–1.1 to 8.7)	26 (7.3)	0.03
Headache							
All related events	Parent/ self	8/581	18.4 (15.3–21.5)	12.5 (8.9–16.0)	5.9 (1.4–10.4)	22 (9.5)	0.55
	Teacher/ staff	4/146	8.0 (2.4–13.6)	12.1 (5.9–18.3)	–4.1 (–11.3 to 3.2)	–21 (13.8)	0.54
Serious side effects	Parent/ self	4/174	6.9 (4.8–9.0)	1.6 (0.0–4.3)	5.3 (1.9–8.7)	18 (1.7)	0.33
Stomach ache							
All related events	Parent/ self	7/290	24.0 (19.0–28.9)	14.9 (8.7–21.1)	9.0 (1.2–16.9)	9 (18.3)	0.44
Serious side effects	Parent/ self	4/201	5.2 (3.4–7.0)	1.2 (0.0–3.6)	4.0 (1.0–6.9)	18 (1.7)	0.78
Drowsiness							
All related events	Parent/ self	4/201	24.3 (16.6–32.0)	14.5 (4.5–24.6)	9.8 (–2.8 to 22.3)	10 (14.3)	0.62
	Teacher/ staff	3/121	12.6 (7.1–18.1)	12.2 (6.6–17.7)	0.5 (–4.7 to 5.6)	367 (8.3)	0.70
Anxiety							
All related events	Parent/ self	7/482	31.1 (24.8–37.5)	38.4 (29.9–46.8)	–7.2 (–17.8 to 3.3)	–9 (56)	0.43
Dizziness							
All related events	Parent/ self	4/383	7.3 (5.5–9.1)	2.2 (0.0–4.6)	5.1 (2.2–8.1)	11 (4.0)	0.31

Note: pts = trial participants, MPH = methylphenidate, NNH = number needed to harm.

*Heterogeneity test.

cal tests. Finally, the outcome dependence of the short-term methylphenidate effect raises questions about the exact nature of the hyperactivity index measure.

It is actually an index of disruptive-externalizing (i.e., hyperactive and antisocial) behaviour,⁴⁵ which is neither necessary nor sufficient to diagnose ADD. The significant hyperactivity index-defined effects were, therefore, expressed in terms of features comorbid to ADD, making it an inappropriate primary outcome in ADD research. Nevertheless, these effects are not surprising given that most of the children in the trials that we analyzed were school-aged males, many of whom had received a diagnosis of ADD with disruptive-externalizing co-features. As a result, whereas 30%–80% of ADD-diagnosed children may also exhibit externalizing disorders,^{13,73} ADD diagnoses that typically exclude these co-features, such as the inattentive type listed in the *Diagnostic and Statistical Manual of Disorders (DSM-IV)*, which may be a more likely diagnosis for females,^{1,10,13,15,21,74} were poorly represented here. This perspective may explain why having some female participants in trials yielded smaller hyperactivity index-defined effect sizes. Our findings are, therefore, not readily generalizable to forms of ADD that exclude disruptive-externalizing co-features, and probably not at all to females.

The clinical implications of the present findings primarily centre on the need to have practitioners recognize that the benefits and risks associated with methylphenidate therapy should be carefully reviewed prior to the start of, and monitored vigilantly during, treatment. In addition, clinicians and policy-makers alike must consider that the received view regarding the short-term, placebo-controlled efficacy and safety of methylphenidate for ADD has likely been based on some cross-section of RCT or non-RCT evidence, unsystematic narrative syntheses or flawed meta-analyses, or some combination of these,^{3,5,10,15,23,24,41} that largely describe males with a restricted definition of ADD. Thus, broad generalizations concerning the usefulness of methylphenidate should probably be avoided.

The research implications suggest the need to undertake a large, long-term trial lasting longer than 14 months to redress the notable methodological and clinical shortcomings of previous efforts (e.g., confounding of dose and dose order, poor washouts) in determining exactly to what extent, in which terms, and for whom (e.g., with all ADD subtypes and females represented), specific courses of short-acting methylphenidate work efficaciously and safely in the short term and the long term. Whatever its design (e.g., placebo-controlled [parallel or crossover⁷⁵] RCT or prospective cohort study), this study will require juggling of both methodological (e.g., control of bias) and ethical considerations (e.g., withholding methylphenidate in a placebo-controlled RCT, which may strongly affect compliance), while also resolving the problem of “diagnosis opacity” whereby a single diagnos-

tic label within a given DSM version refers to different combinations of clinical signs or symptoms (e.g., “select ≥ 6 of these 9 ...”).

To conclude, we found that short-acting methylphenidate was an effective short-term treatment option for children diagnosed with ADD. Yet, this finding may not be robust or completely valid. We also observed that this treatment exhibits a short-term safety profile that requires further investigation. Finally, there is a lack of long-term randomized trial evidence. Collectively, these observations likely reflect a less-than-an-ideal state of affairs given the long history of extensive, and ever increasing, use of methylphenidate for ADD particularly in North America for groups that now include preschoolers and adults.^{9,67,76–79}

Competing interests: None declared.

Contributors: Dr. Schachter was the principal investigator. He contributed to the development of the research question and methodology, planning of the data analysis, the systematic review, data management, derivation of tables, interpretation of results and manuscript preparation. Mr. Pham contributed to the development of the research question and methodology, planning of data analysis, data abstraction and management, statistical analyses, derivation of figures, interpretation of results and revisions of the report. Dr. King contributed to the development of the research question and methodology, planning of data analysis, the systematic review, interpretation of results and revisions of the report. Ms. Langford contributed to data abstraction and management and to revisions of the manuscript. Mr. Moher contributed to the development of the research question and methodology, planning of data analysis, trial quality assessment, interpretation of results and to revisions of the report.

Acknowledgements: We wish to thank the following individuals for their exemplary contributions to this project: Margaret Sampson, for her expertise in searching the literature; Natasha Wiebe, for her help in developing a database; Louise Roy, for her ability to organize events crucial to this project; Dr. Nick Barrowman, for his recommendations and help with matters of a statistical and graphical nature; Dr. Julian Higgins, for his review of an earlier version of the manuscript, including helpful comments concerning the statistical techniques; Drs. Malcolm Rose and Sue Pisterman, for the measurement and clinical information they provided at the inception of this research; and Drs. James Wright and Terry Klassen, for facilitating this research endeavour.

Financial support was provided by the Therapeutics Initiative, University of British Columbia; the British Columbia Ministry for Children and Families; and Cochrane Child Health Field, the Cochrane Collaboration.

References

1. American Psychiatric Association. *Diagnostic and statistical manual of mental disorders (DSM-IV)*. Washington: The Association; 1994.
2. Goldman LS, Genel M, Bezman RJ, Slanetz PJ. Diagnosis and treatment of Attention-Deficit/Hyperactivity Disorder in children and adolescents. Council on Scientific Affairs, American Medical Association. *JAMA* 1998;279:1100–7.
3. Jadao AR, Boyle M, Cunningham C, Kim M, Schachar R. *Treatment of Attention-Deficit/Hyperactivity Disorder*. Prepared for Agency of Healthcare Research and Quality. Rockville (MD): The Agency; 1999.
4. Elia J, Ambrosini P, Rapoport J. Treatment of Attention-Deficit/Hyperactivity Disorder. *N Engl J Med* 1999;340:780–8.
5. Barkley RA, editor. *Attention-deficit hyperactivity disorder. A handbook for diagnosis and treatment*. 2nd ed. New York: The Guilford Press; 1998.
6. Bauchner H. ADHD: a new practice guideline from the American Academy of Pediatrics. *Arch Dis Child* 2000;83:63.
7. Carey WB. What the multimodal treatment study of children with Attention-Deficit/Hyperactivity Disorder did and did not say about the use of methylphenidate for attention deficits. *Pediatrics* 2000;105:863–4.
8. Hartmann T. Whose order is being disordered by ADHD? *Tikkun* 1999;14:17–21.
9. Zito JM, Safer DJ, dosReis S, Gardner JF, Boles M, Lynch F. Trends in the prescribing of psychotropic medications to preschoolers. *JAMA* 2000;283:1025–30.
10. Dulcan M. Practice parameters for the assessment and treatment of children,

- adolescents and adults with attention-deficit hyperactivity disorder. *J Am Acad Child Adolesc Psychiatry* 1997;36(Suppl):85S-121S.
11. Tannock R. Attention deficit hyperactivity disorder: advances in cognitive, neurobiological, and genetic research. *J Child Psychol Psychiatry* 1998;39:65-99.
 12. Shaywitz SE, Shaywitz BA. Attention deficit disorder: diagnosis and role of Ritalin in management. In: Greenhill LL, Osman BB, editors. *Ritalin: theory and patient management*. New York: MA Liebert Publishers; 1991.
 13. Green M, Wong M, Atkins D. *Diagnosis of attention deficit/hyperactivity disorder: technical review 3*. Rockville (MD): Agency for Health Care Policy and Research; 1999. Publication no.: 99-0050.
 14. Jureidini J. Some reasons for concern about attention deficit hyperactivity disorder. *J Paediatr Child Health* 1996;32:201-3.
 15. Schachar R, Tannock R, Cunningham C. Treatment. In: Sandberg S, editor. *Hyperactivity disorders of childhood*. Cambridge: Cambridge University Press; 1996.
 16. Kewley GD. Personal paper: ADHD is underdiagnosed and undertreated in Britain. *BMJ* 1998;316:1594-6.
 17. Swanson JM, Sergeant JA, Taylor E, Sonuga-Barke EJ, Jensen PS, Cantwell DP. Attention-deficit hyperactivity and hyperkinetic disorder. *Lancet* 1998;351:429-33.
 18. DuPaul GJ, Barkley RA, Connor DF. Stimulants. In: Barkley RA, editor. *Attention-deficit hyperactivity disorder. A handbook for diagnosis and treatment*. 2nd ed. New York: The Guilford Press; 1998. p. 510-51.
 19. Greenhill LL. Attention-deficit hyperactivity disorder: the stimulants. *Child Adolesc Psychiatry Clin North Am* 1995;4:123-68.
 20. Greenhill LL, Halperin JM, Abikoff H. Stimulant medications. *J Am Acad Child Adolesc Psychiatry* 1999;38:503-12.
 21. Kupfer DJ, the Consensus Development Panel. National Institutes of Health Consensus Development Conference statement: diagnosis and treatment of Attention-Deficit/Hyperactivity Disorder (ADHD). *J Am Acad Child Adolesc Psychiatry* 2000;39:182-93.
 22. Wigal T, Swanson JM, Regino R, Lerner MA, Soliman I, Steinhoff K, et al. Stimulant medications for the treatment of ADHD: efficacy and limitations. *Ment Retard Dev Disabil Res Rev* 1999;5:215-24.
 23. Weiss G. Attention deficit hyperactivity disorder. In: Lewis M, editor. *Child and adolescent psychiatry: a comprehensive textbook*. 2nd ed. Baltimore: William & Wilkins; 1996.
 24. Cantwell DP. Attention deficit disorder: a review of the past 10 years. *J Am Acad Child Adolesc Psychiatry* 1996; 35:978-87.
 25. Conners CK. *Conners' Rating Scales — revised. Technical manual*. Toronto: Multi-Health Systems; 1997.
 26. Committee on Quality Improvement and Subcommittee on Attention-Deficit/Hyperactivity Disorder. American Academy of Pediatrics' clinical practice guideline: diagnosis and evaluation of the child with attention-deficit/hyperactivity disorder. *Pediatrics* 2000;105:1158-70.
 27. Clinical practice guideline. Diagnosis and evaluation of the child with attention-deficit/hyperactivity disorder. American Academy of Pediatrics. *Pediatrics* 2000;105:1158-70.
 28. Kavale K. The efficacy of stimulant drug treatment for hyperactivity: a meta-analysis. *J Learn Disabil* 1982;15: 280-9.
 29. Kavale K, Forness SR. Hyperactivity and diet treatment: a meta-analysis of the Feingold hypothesis. *J Learn Disabil* 1983;16:324-30.
 30. Ottenbacher KJ, Cooper HM. Drug treatment of hyperactivity in children. *Dev Med Child Neurol* 1983;25:358-66.
 31. Thurber S, Walker CE. Medication and hyperactivity: a meta-analysis. *J Gen Psychol* 1983;108:79-86.
 32. Schachar R, Tannock R. Childhood hyperactivity and psychostimulants: a review of extended treatment studies. *J Child Adolesc Psychopharmacol* 1993;3:81-97.
 33. Wilens TE, Biederman J, Spencer TJ, Prince J. Pharmacotherapy of adult attention deficit/hyperactivity disorder: a review. *J Clin Psychopharmacol* 1995; 15:270-9.
 34. Silva RR, Munoz DM, Aplert M. Carbamazepine use in children and adolescents with features of attention-deficit hyperactivity disorder: a meta-analysis. *J Am Acad Child Adolesc Psychiatry* 1996;35:352-8.
 35. Spencer T, Biederman J, Wilens T, Harding M, O'Donnell D, Griffin S. Pharmacotherapy of attention-deficit hyperactivity disorder across the life. *J Am Acad Child Adolesc Psychiatry* 1996;35:409-32.
 36. Stein MA, Krasowski M, Leventhal BL, Phillips W, Bender BG. Behavioral and cognitive effects of methylxanthines. A meta-analysis of theophylline and caffeine. *Arch Pediatr Adolesc Med* 1996;150:284-8.
 37. Shevell M, Schreiber R. Pemoline-associated hepatic failure: a critical analysis of the literature. *Pediatr Neurol* 1997;16:14-6.
 38. DuPaul GJ, Eckert TL. The effects of school-based interventions for attention deficit hyperactivity disorder: a meta-analysis. *Sch Psychol Rev* 1997;26:5-27.
 39. Crenshaw TM, Kavale KA, Forness SR, Reeve RE. Attention deficit disorder and the efficacy of stimulant medication: a meta-analysis. In: Scruggs TE, Mastropieri MA, editors. *Advances in learning and behavioral disabilities*. Stamford (CT): JAI Press; 1999. p. 135-65.
 40. Gilmore A, Best L, Milne R. *Methylphenidate in children with hyperactivity*. Southampton (UK): Wessex Institute for Health Research and Development; 1998.
 41. Klassen A, Miller A, Raina P, Lee SK, Olsen L. Attention-deficit hyperactivity disorder in children and youth: a quantitative systematic review of the efficacy of different management strategies. *Can J Psychiatry* 1999;44:1007-16.
 42. Jadaar AR, Booker L, Gauld M, Kakuma R, Boyle M, Cunningham CE, et al. The treatment of attention-deficit hyperactivity disorder: an annotated bibliography and critical appraisal of published systematic reviews and meta-analyses. *Can J Psychiatry* 1999;44:1025-35.
 43. Wellbanks L, editor. *Compendium of pharmaceuticals and specialties (CPS)*. 36th edition. Ottawa: Canadian Pharmacists Association; 2001.
 44. Jadaar AR. *Randomised controlled trials*. London: BMJ Books; 1998.
 45. Ullmann RK, Sleator EK, Sprague RL. A change of mind: the Conners Abbreviated Rating Scales reconsidered. *J Abnorm Child Psychol* 1985;13:553-65.
 46. Richters JE, Arnold E, Jensen PS, Abikoff H, Conners K, Greenhill LL, et al. NIMH Collaborative Multisite Multimodal Treatment Study of Children with ADHD: I. Background and rationale. *J Am Acad Child Adolesc Psychiatry* 1995;34:987-1000.
 47. August G, Realmuto G, MacDonald A, Nugent S, Crosby R. Prevalence of ADHD and comorbid disorders among elementary school children screened for disruptive behavior. *J Abnorm Child Psychol* 1996;24:571-93.
 48. American Psychiatric Association. *Diagnostic and statistical manual of mental disorders (DSM-III)*. Washington: The Association; 1980.
 49. Jadaar AR, Moore RA, Carroll D, Jenkinson C, Reynolds DJM, Gavaghan DJ, et al. Assessing the quality of reports of randomized clinical trials: Is blinding necessary? *Control Clin Trials* 1996;17:1-12.
 50. Schulz KF, Chalmers I, Hayes RJ, Altman DG. Empirical evidence of bias: dimensions of methodological quality associated with estimates of treatment effects in controlled trials. *JAMA* 1995;273:408-12.
 51. Berlin JA. Does blinding of readers affect the results of meta-analyses? University of Pennsylvania blinding study group. *Lancet* 1997;350:185-6.
 52. Follman D, Elliott P, Sul I, Cutler J. Variance imputation for overviews of clinical trials with continuous response. *J Clin Epidemiol* 1992;45:769-73.
 53. Earle CC, Pham B, Wells GA. An assessment of methods to combine published survival curves. *Med Decis Making* 2000;20:104-11.
 54. Barkley RA, McMurray MB, Edelbrock CS, Robbins K. Side effects of methylphenidate in children with attention deficit hyperactivity disorder: a systemic, placebo-controlled evaluation. *Pediatrics* 1990;86:184-92.
 55. Hasselblad V. Meta-analysis of multitreatment studies. *Med Decis Making* 1998;18:37-43.
 56. Normand SLT. Tutorial in biostatistics. Meta-analysis: formulating, combining, and reporting. *Stat Med* 1999;18:321-59.
 57. Tacketom CK, Hurlburt Hodding J, Kraus DM. *Pediatric dosage handbook*. 6th edition. Cleveland: Lexi-Comp Inc.; 1999-2000.
 58. American Psychiatric Association. *Diagnostic and statistical manual of mental disorders (DSM-III-R)*. 3rd rev ed. Washington: The Association; 1987.
 59. Begg CB, Berlin JA. Publication bias: a problem in interpreting medical data. *J R Stat Soc* 1988;151:419-63.
 60. Egger M, Davey SG, Schneider M, Minder C. Bias in meta-analysis detected by a simple, graphical test. *BMJ* 1997;315:629-34.
 61. Duval S, Tweedie R. A nonparametric "trim and fill" method of accounting for publication bias in meta-analysis. *J Am Stat Assoc* 2000;95:89-98.
 62. Moher D, Cook DJ, Eastwood S, Olkin I, Rennie D, Stroup DF, for the QUOROM Group. Improving the quality of reports of meta-analyses of randomized controlled trials: the QUOROM statement. *Lancet* 1999;354:1896-900.
 63. Schachar RJ, Tannock R, Cunningham C, Corkum PV. Behavioral, situational, and temporal effects of treatment of ADHD with methylphenidate. *J Am Acad Child Adolesc Psychiatry* 1997;36:754-63.
 64. Law SF, Schachar RJ. Do typical clinical doses of methylphenidate cause tics in children treated for attention-deficit hyperactivity disorder? *J Am Acad Child Adolesc Psychiatry* 1999;38:944-51.
 65. Fischer M, Newby R. Assessment of stimulant response in ADHD children using a refined multimethod clinical protocol. *J Clin Child Psychology* 1991; 20:232-44.
 66. Forness SR, Swanson JM, Cantwell DP, Guthrie D, Sena R. Response to stimulant medication across six measures of school-related performance in children with ADHD and disruptive behavior. *Behav Disord* 1992;18:42-53.
 67. Safer DJ, Zito JM, Fine EM. Increased methylphenidate usage for attention deficit disorder in the 1990s. *Pediatrics* 1996;96:1084-8.
 68. The MTA Cooperative Group. A 14-month randomized clinical trial of treatment strategies for Attention-Deficit/Hyperactivity Disorder. *Arch Gen Psychiatry* 1999;56:1073-86.
 69. Ernst E, Pittler MH. Systematic reviews neglect safety issues. *Arch Intern Med* 2001;161:125-6.
 70. Ioannidis JPA, Contopoulos-Ioannidis DG. Reporting of safety data from randomised trials. *Lancet* 1998;352:1752-3.
 71. Ioannidis JPA, Lau J. Completeness of safety reporting in randomized trials. An evaluation of 7 medical areas. *JAMA* 2001;285:437-43.
 72. Smeeth L, Haines A, Ebrahim S. Numbers needed to treat derived from meta-analyses — sometimes informative, usually misleading. *BMJ* 1999;318: 1548-51.
 73. Biederman J, Newcorn J, Sprich S. Comorbidity of attention deficit hyperactivity disorder with conduct, depressive, anxiety, and other disorders. *Am J Psychiatry* 1991;148:564-77.

74. Gaub M, Carlson CL. Gender differences in ADHD: a meta-analysis and critical review. *J Am Acad Child Adolesc Psychiatry* 1997;36:1036-45.
75. Khan KS, Daya S, Collins JA, Walter SD. Empirical evidence of bias in inferability research: overestimation of treatment effect in crossover trials using pregnancy as the outcome measure. *Fertil Steril* 1996;65:939-45.
76. Coyle JT. Psychotropic drug use in very young children. *JAMA* 2000;283:1059-60.
77. Minde K. The use of psychotropic medications in preschoolers: some recent developments. *Can J Psychiatry* 1998;43:571-5.
78. Rapaport MD, Mullan PB, Alvarez FJ, Eneli IU, Wang J, Gardner JC. Diagnosis of attention deficit disorder and use of psychotropic medication in very young children. *Arch Pediatr Adolesc Med* 1999;153:1039-45.
79. Vimpani GV. Prescribing stimulants for disruptive behaviour disorders: sometimes against the best interest of the child. *J Paediatr Child Health* 1997;33:9-11.

Correspondence to: Dr. Howard M. Schachter, Thomas C. Chalmers Centre for Systematic Reviews, Children's Hospital of Eastern Ontario Research Institute, 401 Smyth Rd., Ottawa ON K1H 8L1; fax 613 738-4869; hschacht@uottawa.ca

Appendix 1: MEDLINE search strategy

1. (attention adj2 deficit adj3 disorder\$.tw.
2. hyperactiv\$.tw.
3. hyperkine\$.tw.
4. psycho?stimul\$.tw.
5. (methylphenidate OR ritalin).tw.
6. exp attention deficit disorder with hyperactivity/
7. exp methylphenidate/
8. exp hyperkinesis/
9. OR/1-3,6,8
10. 5 OR 7
11. randomized controlled trial.pt.
12. clinical trial.pt.
13. controlled clinical trial.pt.
14. crossover\$.tw.
15. cross-over\$.tw.
16. ((singl\$ or doubl\$ or tripl\$ or trebl\$) adj3 (blind\$ or mask \$)).tw.
17. randomi?ed.tw.
18. (random\$ adj3 (select\$ or assign\$ or allocat\$ or sampl\$)).tw.
19. "PLACEBO\$".mp.
20. exp clinical trials/
21. exp case-control studies/
22. exp cohort studies/
23. OR/11-20
24. OR/21-22
25. 23 NOT 24
26. 9 AND 10 AND 25
27. limit 26 to year = 1981-1999
28. limit 27 to (adult < 19 to 44 years > or middle age < 45 to 64 years > or "aged < 65 and over >" or "aged, 80 and over")
29. limit 27 to (infant < 1 to 23 months > or preschool child < 2-5 years > or child < 6 to 12 years >)
30. 28 NOT 29
31. 27 NOT 30

focus on
medical
solutions



Free
to CMA Members

featuring

NEW

MD Consult

Electronic books and journals

Ovid Medline

Fast, precise searching

Medical Librarian Online

Support when you need it

www.cma.ca/osler

1 (800) 457-4205

A benefit of CMA membership



CANADIAN
MEDICAL
ASSOCIATION