

A systematic review and economic model of the effectiveness and cost-effectiveness of methylphenidate, dexamfetamine and atomoxetine for the treatment of attention deficit hyperactivity disorder in children and adolescents

S King,^{1*} S Griffin,² Z Hodges,¹ H Weatherly,²
C Asseburg,² G Richardson,² S Golder,¹ E Taylor,³
M Drummond² and R Riemsma¹



¹ Centre for Reviews and Dissemination, University of York, UK

² Centre for Health Economics, University of York, UK

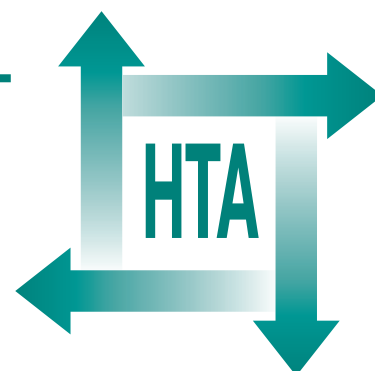
³ Institute of Psychiatry, King's College London, UK

* Corresponding author

Executive summary

Health Technology Assessment 2006; Vol. 10: No. 23

Health Technology Assessment
NHS R&D HTA Programme





Executive summary

Background

Attention deficit hyperactivity disorder (ADHD) (including hyperkinetic disorder) is defined by the 'core' signs of inattention, hyperactivity and impulsivity, and is characterised by an early onset. The estimated prevalence for ADHD in school-aged children varies widely (e.g. 3–7%), being dependent on a number of variables, including the methods of ascertainment, the informants, the population sampled, the diagnostic criteria applied and the sex of the affected individual. Data on prevalence in adolescence and adulthood are limited. The disorder is frequently observed in greater numbers of males than females, with ratios ranging from 2:1 to 9:1 depending on subtype and setting.

There are two generally used diagnostic criteria: the International Statistical Classification of Diseases and Related Health Problems (ICD-10) and *Diagnostic and Statistical Manual of Mental Disorders* (DSM-IV) criteria. The ICD-10 presents details on the diagnosis of hyperkinetic disorders (HKD) and the DSM-IV criteria define ADHD more broadly to include three subtypes: a combined subtype in which all three core signs are present, a predominantly inattentive subtype in which inattention is present but not hyperactivity/impulsivity and a predominantly hyperactive-impulsive subtype in which hyperactivity/impulsivity are present but not inattention. As the ICD-10 criteria are similar to the severe combined type ADHD defined by the DSM-IV criteria, prevalence rates may be higher using the DSM-IV criteria than when diagnosed using the ICD-10 criteria.

Current treatments for ADHD include social, psychological and behavioural interventions in addition to medical management. Medications currently licensed for the treatment of ADHD in the UK include methylphenidate hydrochloride (MPH), dexamfetamine sulphate (DEX) and atomoxetine (ATX), although clinicians sometimes prescribe tricyclic and other antidepressants. MPH is available in immediate-release (Ritalin[®] and Equasym[®]) and extended-release forms [Concerta[®] XL and Equasym XL[®] (a licence application for Equasym XL had been submitted; it has been

specifically developed to provide efficacy across the school day and replace the need for twice daily dosing for children who do not consistently require evening medication)]. They are all indicated in children over 6 years of age, and in adolescents. DEX can be given to children as young as 3 years, whereas ATX (licensed in the UK in May 2004) is indicated in children aged 6 years and above.

Objective

The objective was to assess the clinical and cost-effectiveness of oral MPH, DEX and ATX in children and adolescents (under 18 years of age) diagnosed with ADHD (including hyperkinetic disorder).

Methods

This systematic review incorporated studies from, and built upon, three previous systematic reviews:

- A review conducted by the American Agency for Healthcare Research and Quality (AHRQ) published in 1999 (Jadad and colleagues, 1999).
- A report for the Canadian Coordinating Office for Health Technology Assessment (CCOHTA) (Miller and colleagues, 1999).
- A previous National Institute for Health and Clinical Excellence (NICE) review, which was also based primarily on evidence from the AHRQ report (Lord and Paisley, 2000).

Search strategy

The searches, conducted in July 2004, aimed to retrieve both published and unpublished papers with no language restrictions. A date restriction of 1999 onwards was placed on the methylphenidate searches to update the report produced by Lord and Paisley published in 2000. A date restriction of 1997 onwards was placed on the searches for dexamfetamine to update the AHRQ report (which included a review of this drug). Research on atomoxetine was searched for from 1981 onwards. The search strategy was based on that used in the AHRQ report. ▶

Inclusion/exclusion criteria

To be eligible for inclusion, studies had to be randomised controlled trials (RCTs) of at least 3 weeks' duration (3 weeks per treatment arm in parallel studies and 3 weeks in overall trial length for crossover studies). In addition, systematic reviews were included to examine adverse events data. For the assessment of cost-effectiveness, a broader range of studies was considered.

The studies had to examine MPH, DEX or ATX used alone or in combination with non-drug interventions and be compared with placebo, with one another in head-to-head comparisons or with non-drug interventions. Non-drug interventions included any type of psychological and behavioural strategies (e.g. cognitive behavioural therapy, child or parent training, bibliotherapy) and/or nutritional interventions. Studies that compared MPH, DEX or ATX with other drugs (e.g. Adderall) not licensed in the UK for ADHD were included as long as there was a placebo group. This was applied to both efficacy and adverse events data.

Participants included children and adolescents under 18 years of age diagnosed with ADHD (including hyperkinetic disorder). There was no lower age limitation (although there was no lower age limitation for the report, it is noted that MPH is indicated for children older than 6 years, and ATX is indicated for children aged 6 years and over). The diagnosis must have been made in an explicit way, preferably using either the ICD-10 criteria or the DSM-IV criteria. Studies including participants with conditions other than ADHD (e.g. Tourette's syndrome) were excluded unless they reported separate analyses for patients with ADHD alone.

To be included in the review, trials had to report results on one or more of the following:

- core symptoms (including measures of inattention, hyperactivity, impulsivity)
- quality of life (QoL) (Clinical Global Impression or overall severity indices were used as a proxy of QoL)
- adverse effects (including loss of appetite, insomnia, headache, stomach ache and weight loss).

Studies that only examined tests of psychological function (e.g. the continuous performance test), measures of depression and/or anxiety, or measures of coexistent problems (including poor peer relationships, and conduct/oppositional-

disorder-related outcomes) were not included in the review. Studies that presented results in figures without presenting actual numbers, or only significance values for comparisons, were excluded from the review.

Studies that met the inclusion criteria above, but were only published as abstracts or as conference presentations were not included in the review unless a full paper could be obtained that related to the abstract.

Two reviewers independently screened all titles and abstracts, including economic evaluations, identified in the updated literature search. Full paper manuscripts of any titles/abstracts that were considered relevant by either reviewer were obtained where possible. In addition, full paper copies of relevant studies presented in the NICE, AHRQ and CCOHTA reports were obtained. The full papers were then assessed against the inclusion criteria by one reviewer and checked by another. Any discrepancies were resolved by consensus and, if necessary, a third reviewer was consulted.

Data extraction and quality assessment

The quality of the clinical effectiveness studies was assessed using modified criteria based on CRD Report No. 4. Each study was assessed and data were extracted by one reviewer and independently checked for agreement with a second reviewer. Disagreements were resolved by consensus and, if necessary, a third reviewer was consulted.

Methods of analysis/synthesis

Clinical effectiveness data were reported separately for each drug and by the type of comparison. Data for MPH were also analysed separately based on whether it was administered as an immediate release or extended release formulation. For all drugs, the data were examined by dose. Data for the core outcomes of hyperactivity (using any scale), Clinical Global Impression (as a proxy of QoL) and adverse events were reported. For crossover studies, the mean and standard deviation (SD) for each outcome were data extracted for end of trial data (i.e. baseline data were not considered). Where possible, we aimed to calculate mean difference and standard errors for crossover studies in order to facilitate meta-analysis. However, owing to the lack of information needed to calculate mean differences in many of the studies, this was not possible. For parallel studies, change scores were reported where given, otherwise means and SDs were presented for end of trial data. In

addition, mean differences with 95% confidence intervals were calculated for each study. For adverse events, self-ratings were reported when used, otherwise, parent reports were utilised. Percentages of participants reporting adverse events were used to calculate numbers of events in each treatment arm.

For the cost-effectiveness section of the report, details of each identified published economic evaluation, together with a critical appraisal of its quality, were presented in structured tables.

Handling company submissions

All the clinical effectiveness data included in the company submissions were assessed. Where these met the inclusion criteria they were included in the clinical effectiveness review. All economic evaluations (including accompanying models) included in the company submissions were assessed and detailed assessments of the assumptions underlying the submitted analyses were undertaken.

A new model was developed to assess the cost-effectiveness of the alternative treatments in terms of cost per quality-adjusted life-year. To achieve this, a mixed treatment comparison model was used to estimate the differential mean response rates. Monte Carlo simulation was used to reflect uncertainty in the cost-effectiveness results.

Results

Clinical effectiveness

In the previous systematic reviews (NICE, AHRQ and CCOHTA), 65 studies were identified as potentially relevant to the current systematic review, and full paper copies were ordered. Of these, 40 met the inclusion criteria. In the updated search, a total of 2908 titles and abstracts relating to clinical effectiveness or systematic reviews of adverse events were identified and screened for relevance. Of these, 409 full paper copies were examined in detail and assessed for inclusion. Of these, 20 RCTs and one systematic review met the inclusion criteria. In addition, four commercial-in-confidence papers were included. Overall, this gives a total of 65 papers.

As reported in the previous NICE report, and in the AHRQ and CCOHTA reviews, the plethora of MPH studies suggest that MPH is effective at reducing hyperactivity and improving QoL (as determined by Clinical Global Impression) in

children. It was noted, however, that the majority of studies that evaluated the effectiveness of MPH did not adequately report their study methodology. Hence, the reliability of the study results is not known. There appears to be little evidence supporting a difference in the effectiveness of immediate-release (IR) and extended-release (ER) MPH.

Similarly, DEX also appears to be effective at reducing hyperactivity and improving QoL, although this is based on a small number of studies. Only one study adequately reported the study methodology.

There was consistent evidence that ATX was superior to placebo for hyperactivity and Clinical Global Impression. Studies on ATX more often reported the study methodology well, and the results are likely to be reliable.

Very few studies made direct head-to-head comparisons between the drugs. The previous NICE report stated that there appeared to be little evidence of difference in the effectiveness of MPH and DEX. No recent studies were found in the updated search. Although the studies reported variable results, the one study that reported no statistically significant differences between the two drugs was deemed to be of good quality, whereas the quality of the others was uncertain given the poor reporting of study methodologies.

One study that compared MPH and ATX reported no differences between the drugs for hyperactivity or Clinical Global Impression. This study did not adequately report study methodology, and the results should be interpreted with caution. **[Confidential information removed].**

Few studies were included in the review that examined a non-drug intervention in combination with MPH, DEX or ATX. Generally, the results were variable. The studies were, however, heterogeneous regarding the type of non-drug interventions examined and the scales used to measure outcomes.

Adequate and informative data regarding the potential adverse effects of MPH, DEX and ATX are lacking. Overall, higher dosages of IR-MPH appear to be associated with the occurrence of headache, lost appetite, stomach ache and insomnia compared with placebo. ER-MPH appears to be associated with decreased appetite and increased insomnia. However, a previous systematic review highlighted the need for further

research into somatic complaints, which may be associated with the disorder itself rather than methylphenidate treatment. Similarly, high doses of DEX appear to be associated with decreased appetite and increased sleeping problems. ATX of any dose may impair appetite.

Cost-effectiveness

The review highlighted a number of potential limitations in the existing literature. In particular, the review highlighted limitations in estimating treatment effectiveness and associated utility values. These limitations may stem from a lack of available data. A new economic model was developed for this report. Pooling was limited in the clinical effectiveness review, owing to heterogeneity between trials. However, some degree of pooling is necessary to proceed with an economic model. The issue of heterogeneity was overcome by basing the base case on trials that are more similar in terms of how they measure the outcome of interest. In a series of sensitivity analyses more trials were included by relaxing the criterion of similarity in outcome measurement. Data on resource use associated with ADHD in the UK were lacking, and so the model relies on estimates from experts.

Given the lack of available evidence for statistically significant differences in efficacy between the alternative drugs, the results of the economic model were largely driven by drug cost, in which there are marked differences. The economic evaluation clearly suggests an optimal treatment strategy, that is, DEX first-line, followed by IR-MPH for treatment failures, followed by ATX for repeat treatment failures. If DEX is considered not suitable as a first-line therapy, the optimal strategy is IR-MPH first-line, followed by DEX as second-line and ATX again as third-line. For patients contraindicated to stimulants, ATX is preferred to no treatment. For patients in whom a midday dose of medication is unworkable, ER-MPH is preferred to ATX, and ER-MPH12 appears more cost-effective than ER-MPH8.

The model is not without limitations. As identified in the clinical effectiveness review, the reporting of studies was poor, there are few data to discriminate between the drugs in efficacy or adverse events and there are few data on long-term efficacy and adverse events associated with medical management of ADHD. The data do not allow discrimination between patients with ADHD in terms of ADHD subtype, age,

gender or previous treatment. These caveats must be borne in mind when interpreting the model results.

Conclusions

The main conclusions from this report are as follows:

1. Drug therapy seems to be superior to no drug therapy.
2. No significant differences between the various drugs in terms of efficacy or side effects were found – mainly owing to lack of evidence.
3. The additional benefits from behavioural therapy (in combination with drug therapy) are uncertain.

The main additional feature of the economic model is the consideration of costs. Given the lack of evidence for any differences in effectiveness between the drugs, the model tends to be driven by drug costs, which differ considerably.

Research recommendations

Future trials examining MPH, DEX and ATX should include the assessment of tolerability and safety as a priority. Reporting should be standardised and transparent. Researchers should refer to the CONSORT approach to study design.

Longer term follow-up of individuals participating in trials could further inform policy makers and health professionals. Such data could potentially distinguish between these drugs in a clinically useful way.

In addition, research examining whether somatic complaints are actually related to drug treatment or to the disorder itself would be informative.

Publication

King S, Griffin S, Hodges Z, Weatherly H, Asseburg C, Richardson G, *et al.* A systematic review and economic model of the effectiveness and cost-effectiveness of methylphenidate, dexamfetamine and atomoxetine for the treatment of attention deficit hyperactivity disorder in children and adolescents. *Health Technol Assess* 2006;**10**(23).

NHS R&D HTA Programme

The research findings from the NHS R&D Health Technology Assessment (HTA) Programme directly influence key decision-making bodies such as the National Institute for Health and Clinical Excellence (NICE) and the National Screening Committee (NSC) who rely on HTA outputs to help raise standards of care. HTA findings also help to improve the quality of the service in the NHS indirectly in that they form a key component of the 'National Knowledge Service' that is being developed to improve the evidence of clinical practice throughout the NHS.

The HTA Programme was set up in 1993. Its role is to ensure that high-quality research information on the costs, effectiveness and broader impact of health technologies is produced in the most efficient way for those who use, manage and provide care in the NHS. 'Health technologies' are broadly defined to include all interventions used to promote health, prevent and treat disease, and improve rehabilitation and long-term care, rather than settings of care.

The HTA Programme commissions research only on topics where it has identified key gaps in the evidence needed by the NHS. Suggestions for topics are actively sought from people working in the NHS, the public, service-users groups and professional bodies such as Royal Colleges and NHS Trusts.

Research suggestions are carefully considered by panels of independent experts (including service users) whose advice results in a ranked list of recommended research priorities. The HTA Programme then commissions the research team best suited to undertake the work, in the manner most appropriate to find the relevant answers. Some projects may take only months, others need several years to answer the research questions adequately. They may involve synthesising existing evidence or conducting a trial to produce new evidence where none currently exists.

Additionally, through its Technology Assessment Report (TAR) call-off contract, the HTA Programme is able to commission bespoke reports, principally for NICE, but also for other policy customers, such as a National Clinical Director. TARs bring together evidence on key aspects of the use of specific technologies and usually have to be completed within a short time period.

Criteria for inclusion in the HTA monograph series

Reports are published in the HTA monograph series if (1) they have resulted from work commissioned for the HTA Programme, and (2) they are of a sufficiently high scientific quality as assessed by the referees and editors.

Reviews in *Health Technology Assessment* are termed 'systematic' when the account of the search, appraisal and synthesis methods (to minimise biases and random errors) would, in theory, permit the replication of the review by others.

The research reported in this monograph was commissioned and funded by the HTA Programme on behalf of NICE as project number 03/33/01. The protocol was agreed in July 2004. The assessment report began editorial review in May 2005 and was accepted for publication in July 2005. The authors have been wholly responsible for all data collection, analysis and interpretation, and for writing up their work. The HTA editors and publisher have tried to ensure the accuracy of the authors' report and would like to thank the referees for their constructive comments on the draft document. However, they do not accept liability for damages or losses arising from material published in this report.

The views expressed in this publication are those of the authors and not necessarily those of the HTA Programme, NICE or the Department of Health.

Editor-in-Chief: Professor Tom Walley
Series Editors: Dr Aileen Clarke, Dr Peter Davidson, Dr Chris Hyde,
Dr John Powell, Dr Rob Riemsma and Dr Ken Stein
Managing Editors: Sally Bailey and Sarah Llewellyn Lloyd

ISSN 1366-5278

© Queen's Printer and Controller of HMSO 2006

This monograph may be freely reproduced for the purposes of private research and study and may be included in professional journals provided that suitable acknowledgement is made and the reproduction is not associated with any form of advertising.

Applications for commercial reproduction should be addressed to NCCHTA, Mailpoint 728, Boldrewood, University of Southampton, Southampton, SO16 7PX, UK.

Published by Gray Publishing, Tunbridge Wells, Kent, on behalf of NCCHTA.
Printed on acid-free paper in the UK by St Edmundsbury Press Ltd, Bury St Edmunds, Suffolk.