MONASH University
Business and Economics

West Heidelberg, Victoria,
September 14, 2006

"Has NICE Got It Right?"

Health Technology Assessments (HTAs)
by the National Institute for Health and Clinical Excellence

Michael Schlander

University of Applied Economic Sciences Ludwigshafen & University of Heidelberg[1]
Institute for Innovation & Valuation in Health Care (INNOVAL[HC]), Germany

HAS NICE GOT IT RIGHT?

An old German saying ...

**"Wer am Wege baut,
hat viele Meister"[1]**

"A house built by
the wayside
is either too high
or too low."

[1]**Martin Luther (1530)**

OUTLINE

Overview

- **HTAs by the National Institute for Health and Clinical Excellence (NICE)**
  - Praise (not only) for Process
- **Is NICE Infallible?**
  - Research Question and Approach
  - Motivation, Strengths, and Limitations
  - NICE Technology Appraisal No. 98
- **Has NICE Got It Right?**
  - NICE Accountability for Reasonableness
  - Some Suggested Underlying Issues
- **Some Implications**

HEALTH TECHNOLOGY ASSESSMENTS (HTAs) BY NICE

- **Praise for Approach**
- **Praise for Process**
- Research Question
- Motivation and Limitations
- Qualitative Case Study: TA No. 98
- Accountability for Reasonableness
- Some Suggested Underlying Issues
- Some Implications

NICE

National Institute for Health and Clinical Excellence

- Three (distinct) "**Centres of Excellence**":
- **Centre for Public Health Excellence**
  - Public health guidance
    on the promotion of good health and the prevention of ill health
- **Centre for Health Technology Evaluation**
  - Technology appraisals (recommendations on the use of new and existing medicines and treatments within the NHS)
  - Interventional procedure guidance (evaluates the safety and efficacy of such procedures where they are used for diagnosis or treatment)
- **Centre for Clinical Practice**
  - Clinical guidelines
    (recommendations, based on the best available evidence, on the appropriate treatment and care of people with specific diseases and conditions)

[1]http://www.nice.org.uk;
last accessed September 13, 2006

Technology Appraisal of Methylphenidate, Dexamphetamine and Atomoxetine (NICE 2006)

NICE Technology Appraisal Process

- **Three (to four) phases**
  - **Scoping**
  - **Assessment**
  - **Appraisal**
  - **Appeal** (if lodged by one or more consultees)
- **Frequently acclaimed features**
  - NICE objective of appraising the evidence in a way that is **"objective, unbiased, and methodologically sound"**[1]
  - An appraisal process that can be described as being **"inclusive, consultative, transparent"**[1]

[1]C. Longson, ISPOR Annual Meeting,
Arlington, VA, May 20, 2001

Technology Appraisal of Methylphenidate, Dexamphetamine and Atomoxetine (NICE 2006)

NICE Technology Appraisal Process

1. Scoping

- DoH develops *remit*; NICE develops draft *scope*
- Ministers select topics suitable for referral
- Consultation on draft remit and draft scope with consultees, commentators, & Assessment Group
- Scoping workshop and invitation by NICE to stakeholders to discuss the appraisal scope
- Final remit produced by DoH and WAG; final scope produced by NICE
- Ministers make *final decision on referral*
- NICE issues *final remit and scope*

HTAs BY NICE

Technology Appraisal of Methylphenidate, Dexamphetamine and Atomoxetine (NICE 2006)

NICE Technology Appraisal Process

2. Assessment

- Assessment Group (AG) formally commissioned to prepare Assessment Report (AR) based on its *assessment protocol*
- Submissions by manufacturers and sponsors
- Preparation of *Assessment Report* (AR) ("reference case" and template defined by NICE, content and quality responsibility of its authors)
- AR sent to consultees and commentators, with confidential information removed
- *Economic model* considered confidential (intellectual property of assessment groups)

NICE Technology Appraisal Process

3. Appraisal

- Appraisal Committee (AC, a standing advisory committee of NICE) considers *Evaluation Report* (including AR) and comments from consultees on AR (including the AG's response to comments, if any)
- AC prepares *Appraisal Consultation Document* (**ACD**); following instructions by the AC, a NICE project team drafts the ACD
- ACD distributed to consultees and commentators
- AC reviews comments on ACD and prepares *Final Appraisal Determination* (**FAD**) document

HTAs BY NICE

Technology Appraisal of Methylphenidate, Dexamphetamine and Atomoxetine (NICE 2006)

NICE Technology Appraisal Process

4. Appeal

(optional)

- FAD distributed and published as ***NICE Guidance*** unless one or more *consultees* lodge an *appeal* within 15 working days from receipt of the FAD
- Appeal is permissible on the following grounds:
  - NICE has failed to act *fairly* and in accordance with its published *procedures*,
  - the FAD is *perverse* in the light of the evidence submitted, with "perverse" meaning that the FAD is "*obviously and unarguably wrong*, in defiance of logic, or so absurd that no reasonable Appraisal Committee could have reached such conclusions", or
  - NICE has *exceeded its powers*.
- *New evidence* or simply *disagreement* with a FAD will not be accepted in this last stage of the appraisal process.
- Nor is it possible to reopen arguments and issues on which a determination by NICE has been reached.

© Michael Schlander – Seminar at Monash University, September 14, 2006

NICE Standard: The Reference Case[1]

¬ Problem definition	¬ Scope from NICE
¬ Comparator	¬ Routine therapies in NHS
¬ Evidence on outcomes	¬ Systematic review
¬ Economic evaluation	¬ Cost-effectiveness analysis
¬ Perspective on outcomes	¬ All health effects on individuals
¬ Perspective on costs	¬ National Health Service
¬ Discount rate	¬ 3.5% p.a. on costs and health effects
¬ Addressing uncertainty	¬ Probabilistic sensitivity analysis
¬ Measure of health benefits	¬ Quality adjusted life-years
¬ Source of preference data	¬ Representative sample of the public
¬ Health state valuation method	¬ Choice-based method - e.g. SG or TTO
¬ Description of health states for calculating QALYs	¬ Using a standardized and validated generic instrument
¬ Equity position	¬ Each additional QALY has equal value

Reference Case Analysis[1]

- **Major changes that NICE introduced in April 2004 included:**
  - Explicit 'Reference Case'
  - No more differential discounting
  - Use of probabilistic sensitivity analysis to address decision uncertainty
  - Explicit consideration of subgroup analysis

IS NICE INFALLIBLE?

- Praise for Approach
- Praise for Process
- **Research Question**
- **Motivation and Limitations**
- **Introduction to Case Study: TA No. 98**
- Accountability for Reasonableness
- Some Suggested Underlying Issues
- Some Implications

INTRODUCTION

The Disorder

Attention-Deficit/Hyperactivity Disorder (ADHD)

Source: www.sagen.at

ADHD – A Challenge for Analysis

- International variation in preferred diagnostic criteria
- International variation in standards of care
- Co-existing disorders (comorbidity)
- Increasing diagnostic prevalence
- Variety of instruments to measure clinical outcomes
- Controversial validity of QALYs in pediatric populations
- Changing therapeutic landscape
- New medications with improved dosing schedules

Clinical Characteristics[1]

- **Inability**
  - To marshal and sustain attention
  - To modulate activity level
  - To moderate impulsive actions
- Resulting in **maladaptive behaviors inconsistent with age and developmental level**
- **Three Types** (DSM-IV-R)
  - Combined inattentive, hyperactive, and impulsive (~80% of patients)
  - Predominantly inattentive (~10-15% of patients)
  - Predominantly hyperactive and impulsive (~5% of patients)

According to DSM-IV, the diagnosis requires evidence of inattention or hyperactivity and impulsivity or both; symptoms that cause impairment – must be present before 7 years of age – must be present in two or more settings (e.g., home, school, or work) – do not occur exclusively during the course of a pervasive developmental disorder, schizophrenia, or another psychotic disorder – are not better accounted for by another mental disorder (e.g., a mood disorder or an anxiety disorder)

[1]cf. M.D. Rappley, NEJM 2005

Diagnostic Criteria

"ADHD" (DSM-IV)

- **Inattention**
  - ≥ 6/9 symptoms

  and / or
- **Hyperactivity and impulsivity**
  - ≥ 6/9 symptoms
- **Symptoms** causing *impairment*
  - Have persisted for ≥ 6 months
  - Are present before 7 years of age
  - Are "pervasive", i.e., present in ≥ 2 settings (e.g., home, school, work)
  - Are not better accounted for by another mental disorder

"HK[C]D" (ICD-10)

- **Inattention** (≥ 6/9 symptoms)

  and
- **Hyperactivity** (≥ 3/5 symptoms)

  and
- **Impulsivity** (≥ 1/4 symptoms)
- **Symptoms** criteria like DSM-IV[1]
- Hyperkinetic Disorder:
  - If criteria above are met (-> F90.0)
- Hyperkinetic Conduct Disorder:
  - If additional symptoms of conduct disorder are present (-> F90.1)

[1]Note that ICD-10 criteria are also stricter than DSM-IV in terms of (a) pervasiveness requirements (demanding more than "impairment") and (b) exclusion of co-existing conditions.

Evidence-Based Treatment[1]

- **Pharmacologic Treatment**
  - Psychostimulants
    - > 250 studies (mostly cross-over trials)
    - N > 5,000)
  - Noradrenergic compounds
- **Behavior Modification**
  - ~48 classroom studies (N > 900)
  - ~80 parent training studies (N > 5,000)
- **The combination of pharmacologic treatment and behavior modification**
  - 25 studies (N > 5,000)

[1]From W.E. Pelham 2005

INTRODUCTION

Acquisition costs of important drugs licensed for treatment of ADHD

Prescription Drug Spending: Acquisition Costs[1]

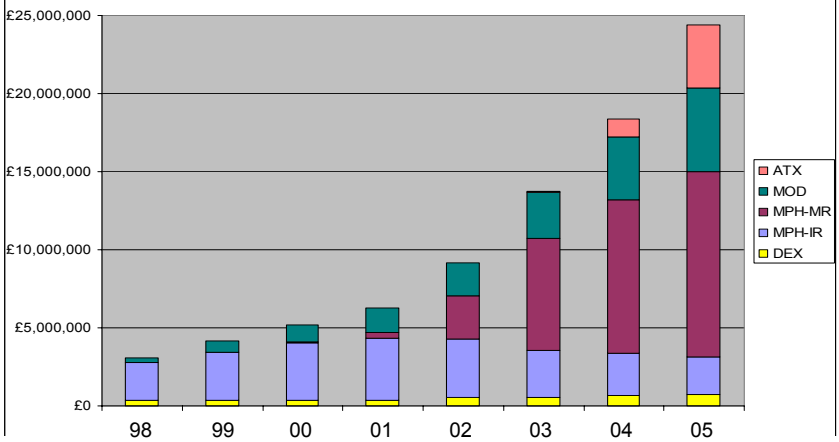
Trade Name	Active Ingredient	Cost / Daily Dose[3]	Assumed *Average* Daily Dose[2]	Daily Dosage Schedule[2]
Dexedrine[R]	Dexamphetamine sulphate	**£ 0.42**	20mg/d	2 times
Ritalin[R]	Methylphenidate hydrochloride	**£ 0.56**	30mg/d	3 times
Equasym[R]	Methylphenidate hydrochloride	**£ 0.56**	30mg	3 times
MPH Generics	Methylphenidate hydrochloride	**<£ 0.56**	30mg	3 times
Equasym[R] XL	Methylphenidate hydrochloride	**£ 1.17**	30mg	1 time
Concerta[R] XL	Methylphenidate hydrochloride	**£ 1.23**	36mg	1 time
Strattera[R]	Atomoxetine hydrochloride	**£ 1.95** (to £ 3.80)	Irrelevant due to flat pricing	1(to 2) times

[1]2005; data sources: UK: British National Formulary (BNF), March 2005 (Equasym XL: September 2005);
[2]assumptions underlying cost data provided here, not to be construed as treatment recommendations since ADHD medication require individual titration;
[3]note that individual doses and hence costs may vary.

INTRODUCTION

The Example of Prescription Drug Spending

ADHD-Related Expenditures (NHS England)[1]



[1]Expenditures by category p.a.; DEX: dexamphetamine (Dexedrine[R] and others); MPH: methylphenidate; IR: immediate-release formulations (Ritalin[R] and generics); MR: modified-release formulations (Concerta[R] XL, Equasym[R] XL; Ritalin[R] SR imports); MOD: modafinil (Provigil[R], licensed for daytime sleepiness); ATX: atomoxetine (Strattera[R]); PEM: pemoline (Volital[R], before 2002 only, not shown due to small volume); data source: NHS Prescription Cost Analysis 1999-2006.

Explaining the profound increase in expected prescription drug spending

Reasons for Increased Spending on ADHD Treatment

1. Growing awareness (education & promotional efforts by industry)
   - ¬ **ADHD being diagnosed more frequently (and earlier)**
2. Growing acceptance of pharmacotherapy
   - ¬ **More patients receiving pharmacotherapy**
3. Increasing intensity of pharmacotherapy
   - ¬ **More prescriptions per diagnosed and treated patient**
4. Improved therapeutic options
   - ¬ **Higher unit cost per defined dose**

¬ These factors combined exert a **multiplicative effect**, leading to the expectation of a pronounced increase of drug expenditures.

¬ Other cost components (including, but not limited to, diagnostic procedures and cognitive-behavioral therapy) are likely to increase as well.

Schlander (2004, 2006)

A broader perspective

ADHD: Burden of Disease (Overview)[1]

¬ **Health Care System**
- ¬ Increased health care utilization and direct medical costs (reported to be comparable to children with asthma); including emergency room visits
- ¬ Increased risk of substance abuse disorders (including earlier onset and lower probability to quit in adulthood)
- ¬ Increased risks of bike and more motor vehicle accidents

¬ **School and Occupation**
- ¬ Many expelled; increased drop-out rates; impaired educational outcomes and lower occupational status

¬ **Family and Employers**
- ¬ Parental divorce (or separation) rates increased; sibling fights
- ¬ Parental absenteeism and productivity

¬ **Society**
- ¬ Criminal behavior; justice and legal system costs, substance abuse disorders

[1]multiple references

IS NICE INFALLIBLE?

- Praise for Approach
- Praise for Process
- Research Question
- Motivation and Limitations
- **Qualitative Case Study: TA No. 98**
- Accountability for Reasonableness
- Some Suggested Underlying Issues
- Some Implications

REVIEW

Technology Appraisal of Methylphenidate, Dexamphetamine and Atomoxetine (NICE 2006)

Assessment Scope[1]

- **Comparators**
  - Include placebo and usual care.
- **Outcomes**
  - Should include the incidence and severity of **core symptoms**, **problem behaviors**, **educational performance**, measures of **depression and / or anxiety**, measures of **conduct / oppositional-disorder**-related outcomes, **adverse events**, and **quality of life**.
- **Consideration**
  - Should be given to the impact of co-morbid disorders, quality of life of other family members, and the optimal duration of treatment...

Clinical Guidelines Remit[2]

- **Management of ADHD**
  - "To prepare a guideline ... on the effectiveness of methylphenidate and other pharmacological **and psychological interventions** in combination or separately for the treatment of ADHD"
  - "The guideline should apply to the treatment of children, young people **and adults** where evidence of treatment effectiveness is available."
- The guideline development process will be led by the **National Collaborating Centre for Mental Health** and is broader in scope[3] than the technology appraisal, intended to cover "**the full range of care routinely made available by the NHS**".

[1]NICE (2003); DoH (2004) and NICE (2004)

[3]Draft Scope, NICE (2006)

ADHD: NICE Assessment Protocol[1]

Outcomes

Data on the following outcome measures (as reported by the participant, parent, teacher or clinician) will be included:

- *incidence and severity of **core symptoms**;*
- *incidence and severity of **coexistent problems** including poor peer relationships, and conduct/oppositional-disorder-related outcomes;*
- ***educational performance**;*
- *measures of **depression and/or anxiety**;*
- ***adverse effects** (including substance abuse);*
- ***quality of life** (including global social adjustment).*

*Studies that have used **parent and teacher rating scales of hyperactivity** will be assessed in the first instance. In addition, physician ratings of clinical global impression will be examined. Alternatively, we will examine any of the outcomes listed above. If the evidence allows, consideration will be given to the use of pharmacological treatments in the presence of co-morbid disorders, the effect of treatments on quality of life of other members of the family, and the optimal duration of treatment before attempting discontinuation and reassessment.*

[1]assessment protocol; King et al., 2004

ADHD Assessment: Search Strategy[1]

- ***Conference proceedings***
- ***Gray literature***
- ***Randomized controlled clinical trials***
  - *of at least three weeks duration*
- ***Full economic evaluations that compare at least two options and consider both costs and consequences, including***
  - *cost-effectiveness,*
  - *cost-minimization,*
  - *cost-utility*
  - *cost-benefit analysis*
- ***"Full paper manuscripts of any titles / abstracts that may be relevant will be obtained where possible"***

[1]assessment protocol; King et al., 2004

NICE 2006: Clinical Evidence Informing Economic Model

Evidence base

Literature search: 2,908 publication titles identified and screened (AR, p. 52)
AHRQ Review (Jadad et al., Nov. 1999): 78 trials (77 RCTs) selected
CCOHTA Review (Miller et al., Dec. 1998): 26 trials selected (n~1,000)
Schachar et al. (2002): 14 trials (≥ 12 weeks) selected (n=1,379)
MTA Cooperative Study Group (1999): 4 groups, 2 years, n=579
Klein et al. (2004), **Abikoff** et al. (2004): 3 groups, 2 years, n=103

Filter 1

RCTs examining MPH, DEX, or ATX,
alone or in combination, with or without NDT;
patients age <18y; "≥3 weeks treatment duration";
reporting core symptoms, quality of life,
adverse effects, or educational performance

Effectiveness review
Focus on hyperactivity ratings

64 randomized clinical studies (n~7,000)
plus
NIMH MTA Study (n= 435 out of n=579)

(>1/3 of these studies were short-term
[<3 weeks *treatment* duration] cross-over trials)

REVIEW

Technology Appraisal of Methylphenidate, Dexamphetamine and Atomoxetine (NICE 2006)

NICE 2006: Clinical Evidence Informing Economic Model

Effectiveness review
Focus on hyperactivity ratings

64 randomized clinical studies (n~7,000)
(hereof, >1/3 short-term cross-over trials
with <3 weeks observation period per treatment arm)
plus
NIMH MTA Study (n= 435 out of n=579)

Filter 2

Availability
of
CGI-I scores
(subscale)

Economic model
Focus on CGI-I scores

5 clinical studies (n= 1,926), treatment duration 3–8 weeks,
hereof 1 study with n=1,323 (Kemner et al., 2004)
and 1 study "CIC";
plus
1 cross-over study previously excluded, n=32 (Sharp et al., 1999)

NICE 2006: Clinical Evidence Informing Economic Model

Economic model
Focus on CGI-I scores

5 clinical studies (n= 1,926), treatment duration 3–8 weeks,
hereof 1 study with n=1,323 (Kemner et al., 2004)
and 1 study "CIC";
plus
1 cross-over study previously excluded, n=32 (Sharp et al., 1999)

Secondary extensions

Availability
of **CGI-S**
or **ADHD-RS**
or **SNAP-IV**

Extended economic model
Focus on "response rates"
defined by four different scales
(or "subscales" in the case of CGI-I and CGI-S)

13 clinical studies (n≥2,768); 4 studies "CIC",
one "CIC" study could not be identified
plus
3 arms of NIMH MTA Study (n=435 out of n=579)

NICE 2006: Economic Model[1]

Studies used in the base case analysis:
Response defined as score of 1 or 2 on the CGI-I subscale

Trial	Treatment	Responders (%)	Number in group
Sharp *et al.* 1999*[145]	IR-MPH	26 (81)	32
	DEX	27 (84)	32
	Placebo	5 (16)	32
Greenhill *et al.* 2002[65]	ER-MPH8	125 (81)	154
	Placebo	78 (50)	156
Kemner *et al.* 2004[96]	ER-MPH12	583 (69)	850
	ATX	250 (53)	473
Steele *et al.* 2004[85]	ER-MPH12	58 (83)	70
	IR-MPH	45 (62)	73
Pliszka *et al.*2000[48]	IR-MPH	13 (65)	20
	Adderall	18 (90)	20
	Placebo	5 (28)	18
Klein *et al.* 1997[92]	IR-MPH + BT	28 (97)	29
	IR-MPH	23 (79)	29
	Placebo + BT	14 (50)	28

* not currently reviewed in chapter 4

[1]Assessment report, p. 226; King et al., 2004

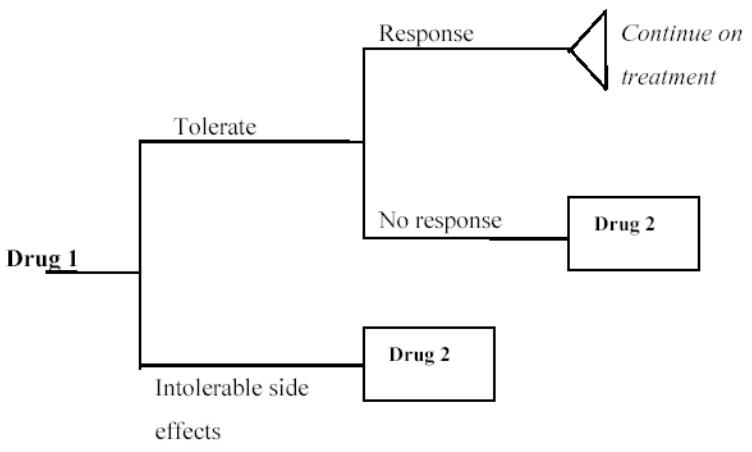
NICE 2006: Economic Model[1]

Studies used in the base case analysis

Study	Comp-arators	Study Design	Study Patients	Endpoints used	Notes
Sharp et al., 1999	MPH-IR DEX Plac.	RCT double-blind 3x crossover (3x3 weeks)	n=32 (girls only)	CGI-I	Excluded from effectiveness review (for "inadequate data presentation"); no data provided in AR; inclusion "initially" based on DSM-IIIR, "later" DSM-IV, combined type
Greenhill et al., 2002 (32 sites)	MPH-MR08 Plac.	RCT PG (1:1) double-blind 3 weeks	n=314 (82% male)	CGI-I CGI-S	Primary endpoint: Conners' Teacher Global Index; study listed among MPH-ER medium dose group in AR (average dose 40.7mg/d)
Kemner et al., 2004 ("multiple sites")	ATX MPH-MR12	RCT PG (2:1) open-label 3 weeks	n=1,323 (74% male)	CGI-I ADHS-RS	"CIC" (no data provided in AR);primary endpoint: ADHD-RS improvement (change in mean score): MPH-MR12 superior to ATX (but included also patients with prior stimulant treatment)
Steele et al., 2004, 2006	MPH-IR MPH-MR12	RCT, PG (1:1) open-label, "real-world" design; 8 weeks	n=145 (83% male)	CGI-I CGI-S? SNAP-IV	"CIC" (no data provided in AR); primary endpoint: SNAP-IV (18/26 items, parent ratings); real-world effectiveness trial; MPH-MR12 superior to MPH-IR
Pliszka et al., 2000 ;	MPH-IR MAS Plac.	RCT double-blind PG (1:1:1) 3 weeks	n=58 (% males ?)	CGI-I	Primary endpoint: IOWA Conners' ratings
Klein and Abikoff, 1997	MPH-IR (w/ and w/o NDT) Plac.	RCT double-blind PG (1:1:1) 8 weeks	n=86 (94% male)	CGI-I	Primary endpoints: CTRS, CPRS; multiple further assessments

[1]Assessment report, King et al., 2004

NICE 2006: Economic Model Structure[1]



[1]Assessment report, p. 223; King et al., 2004

NICE 2006: Economic Model[1]

Treatment sequences compared in economic model

Treatment sequences	1	2	3	4	5	6
1st line	MPH	MPH	ATX	ATX	DEX	DEX
2nd line	ATX	DEX	MPH	DEX	MPH	ATX
3rd line	DEX	ATX	DEX	MPH	ATX	MPH
4th line	No trt	No trt	No trt	No trt	No trt	No trt
	x3 to represent each formulation of MPH = 18 x2 to include combination therapy = 36 + no treatment = 37					

MPH = methylphenidate; ATX = atomoxetine; DEX = dexamfetamine; No trt = no treatment

[1]Assessment report, p. 222; King et al., 2004

REVIEW

Technology Appraisal of Methylphenidate, Dexamphetamine and Atomoxetine (NICE 2006)

NICE 2006: Economic Model[1]

Response and withdrawal rates used in base case analysis

Response defined as score of 1 or 2 on the CGI-I subscale

Treatment	Response rate (standard deviation)	Withdrawal rate (standard deviation)
Placebo	0.28 (0.04)	0.11 (0.02)
IR-MPH	0.68 (0.30)	0.09 (0.05)
ER-MPH8	0.57 (0.33)	0.08 (0.06)
ER-MPH12	0.75 (0.32)	0.12 (0.04)
ATX	0.67 (0.37)	0.11 (0.06)
DEX	0.75 (0.32)	0.02 (0.05)

[1]Assessment report, p. 236; King et al., 2004

NICE 2006: Data used in calculating withdrawal rates[1]

Trial	Treatment	Withdrawals (%)	Number in group
Sharp *et al.* 1999*[145]	IR-MPH DEX Placebo	1 (3) 0 (0) 0 (0)	32 32 32
Greenhill *et al.* 2002[65]	ER-MPH8 Placebo	20 (13) 32 (20)	158 163
Kemner *et al.* 2004[96]	ER-MPH12 ATX	41 (5) 26 (5)	850 473
Steele *et al.* 2004[85]	ER-MPH12 IR-MPH	12 (16) 12 (16)	73 74
Pliszka *et al.*2000[48]	IR-MPH Adderall Placebo	1 (5) 2 (10) 2 (11)	20 20 18
Klein *et al.* 1997[92]	IR-MPH + BT IR-MPH Placebo + BT	0 (0) 1 (3) 2 (7)	29 31 29
Kelsey *et al.* 2004[82]	ATX Placebo	26 (20) 17 (27)	133 64
Michelson *et al.* 2002[74]	ATX Placebo	12 (14) 11 (13)	85 86
Weiss *et al.* 2004[87]	ATX Placebo	17 (17) 4 (8)	101 52
Spencer *et al.* 2002[77]	ATX Placebo	8 (6) 7 (6)	129 124

* not currently reviewed in chapter 4

[1]Assessment report, p. 231; King et al., 2004

Technology Appraisal of Methylphenidate, Dexamphetamine and Atomoxetine (NICE 2006)

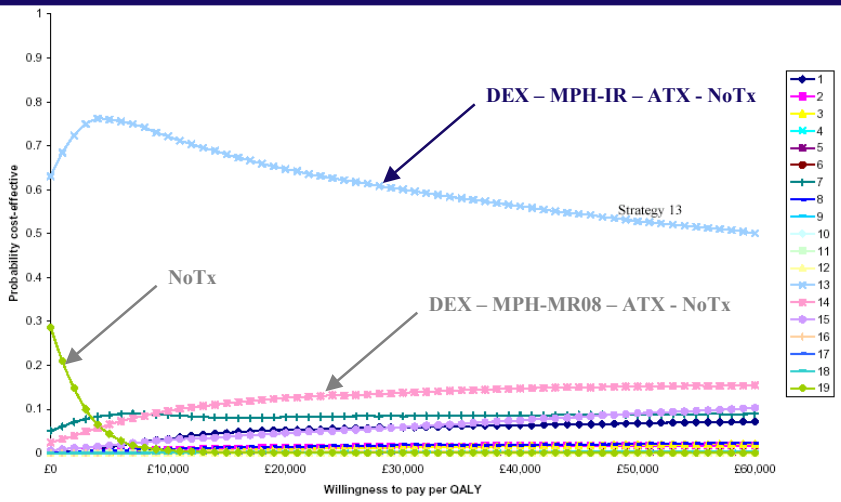
NICE 2006: Base Case Results of the Economic Model[1]

Strategy	Order of treatments received	Cost	QALYs
1	IR-MPH – ATX – DEX - No treatment	£1,233	0.8279
2	ER-MPH8 – ATX – DEX - No treatment	£1,470	0.8273
3	ER-MPH12 – ATX – DEX - No treatment	£1,479	0.8278
4	ATX – IR-MPH – DEX – No treatment	£1,480	0.8278
5	ATX – ER-MPH8 – DEX – No treatment	£1,550	0.8277
6	ATX – ER-MPH12 – DEX – No treatment	£1,563	0.8274
7	IR-MPH – DEX - ATX - No treatment	£1,140	0.8283
8	ER-MPH8 – DEX - ATX - No treatment	£1,336	0.8277
9	ER-MPH12 – DEX - ATX - No treatment	£1,410	0.8284
10	ATX – DEX – IR-MPH– No treatment	£1,466	0.8281
11	ATX – DEX – ER-MPH8 –No treatment	£1,485	0.8281
12	ATX – DEX – ER-MPH12– No treatment	£1,488	0.8278
13	**DEX – IR-MPH – ATX – No treatment**	**£1,098**	**0.8289**
14	DEX – ER-MPH8 – ATX – No treatment	£1,157	0.8287
15	DEX – ER-MPH12 – ATX – No treatment	£1,159	0.8287
16	DEX – ATX – IR-MPH– No treatment	£1,158	0.8288
17	DEX– ATX – ER-MPH8 –No treatment	£1,177	0.8288
18	DEX– ATX – ER-MPH12 –No treatment	£1,180	0.8285
19	No treatment	£1,223	0.7727

[1]Assessment report, p. 237; King et al., 2004

Technology Appraisal of Methylphenidate, Dexamphetamine and Atomoxetine (NICE 2006)

NICE 2006: Base Case Cost-Effectiveness Acceptability Curves[1]



[1]Assessment report, p. 238; King et al., 2004

Technology Appraisal of Methylphenidat… …etamine … …etine (NICE 2006)

NICE 2006: Base Case … of the Econo… …del[1]

Strategy	Or…	Cost	QALYs
1		…23	0.8279
2			0.8273
3			0.8278
			0.8278
			0.8277
			0.8274
			0.8283
			0.8277
			0.8284
			0.8281
			0.8281
			0.8278
			0.8289
			0.8287
			0.8287
16			0.8288
17			0.8288
18		…,180	0.8285
19	No treatm…	£1,223	0.7727

How strong is our confidence
in QALY differences in pediatric populations[2]
extending to the third or fourth decimal place ...

... based upon
¬ **CGI-I response rates** (1 or 2 on a scale of 7)
based on short-term studies
(some involving small patient numbers)
¬ **Relative efficacy**
derived from indirect evidence
(mixed treatment comparison;
heterogeneity problems)
¬ **Utility values**
from EQ-5D-based parent proxy-ratings
¬ **Withdrawal rates**

[1]Assessment report, p. 237; King et al., 2004

[2]cf. Griebsch et al., QALYs lack quality in pediatric care – a critical review. *Pediatrics* 2005; 115: 600-614

NICE 2006: Base Case Results of the Economic Model[1]

Strategy	Order of treatments received	Cost	QALYs
1	IR-MPH – ATX – DEX - No treatment	£1,233	0.8279
2	ER-MPH8 – ATX – DEX - No treatment	£1,470	0.8273
3	ER-MPH12 – ATX – DEX - No treatment	£1,479	0.8278
4	ATX – IR-MPH – DEX – No treatment	£1,480	0.8278
5	ATX – ER-MPH8 – DEX – No treatment	£1,550	0.8277
6	ATX – ER-MPH12 – DEX – No treatment	£1,563	0.8274
7	IR-MPH – DEX - ATX - No treatment	£1,140	0.8283
8	ER-MPH8 – DEX - ATX - No treatment	£1,336	0.8277
9	ER-MPH12 – DEX - ATX - No treatment	£1,410	0.8284
10	ATX – DEX – IR-MPH – No treatment	£1,466	0.8281
11	ATX – DEX – ER-MPH8 – No treatment	£1,485	0.8281
12	ATX – DEX – ER-MPH12 – No treatment	£1,488	0.8278
13	**DEX – IR-MPH – ATX – No treatment**	**£1,098**	**0.8289**
14	DEX – ER-MPH8 – ATX – No treatment	£1,157	0.8287
15	DEX – ER-MPH12 – ATX – No treatment	£1,159	0.8287
16	DEX – ATX – IR-MPH– No treatment	£1,158	0.8288
17	DEX– ATX – ER-MPH8 –No treatment	£1,177	0.8288
18	DEX– ATX – ER-MPH12 –No treatment	£1,180	0.8285
19	No treatment	£1,223	0.7727

[1]Assessment report, p. 237; King et al., 2004

REVIEW

Technology Appraisal of Methylphenidate, Dexamphetamine and Atomoxetine (NICE 2006)

NICE 2004: Main Conclusions of Assessment

- "**Drug therapy seems to be superior to no drug therapy.**
- **No significant differences** between the various drugs in terms of **efficacy** or side effects were found – mainly due to lack of evidence.
- The additional benefits from **behavioral therapy** (in combination with drug therapy) are uncertain"[1].
- "Given the lack of evidence for any differences in **effectiveness** between the drugs, the [economic] model tends to be driven by drug cost, which differ considerably"[1].
- "For a decision taken now, with current available data, **the results of the economic model clearly identify an optimal treatment strategy**"[2] and "this analysis showed that a [...] strategy of 1st line dexamphetamine, followed by 2nd line methylphenidate immediate-release for treatment failures, followed by 3rd line atomoxetine for repeat treatment failures was optimal."

[1]Assessment report, p. 20; King et al., 2004; [2]AR, p.261

NICE 2006: Appraisal Summary

- Where drug treatment is considered appropriate, methylphenidate, atomoxetine, and dexamphetamine are recommended within their licensed indications.
- There are no significant differences between individual drugs in terms of efficacy or side effects
  – *a conclusion derived as a consequence of paucity of evidence used for assessment:*
- Given the limited data used to inform response and withdrawal rates, it is not possible to distinguish between the different strategies on the grounds of cost-effectiveness.
- If there is a choice of more than one appropriate drug, the product with the lowest cost should be prescribed.

NICE 2006: Appraisal Recommendations

- The decision about choice of intervention should be based on
  - The presence of comorbid conditions (e.g., tic disorders, Tourette's syndrome, epilepsy).
  - The adverse event profile.
  - Compliance issues (e.g., the need to administer a mid-day dose at school, and its associated implications).
  - The individual preferences of the patient and/or parent/guardian.

NICE ADHD Technology Assessment 2006 – A Critique (1)

- **Narrow scope**
  - Excluding psychosocial interventions (and long-term sequelae)
  - Role of diagnostic criteria and coexisting conditions not addressed (though included in scope)
- **Data selection for assessment**
  - Idiosyncratic interpretation and/or violation of search criteria
  - Reliance on CGI-I subscores for primary economic analysis, economic model departing from clinical effectiveness review.
  - Reliance on short-term data (3-8 weeks in primary model) to extrapolate long-term outcomes (one year; extensions up to 12 years)
- **Efficacy versus effectiveness distinction**
  - Compliance differences effectively "assumed away" for modeling, with potential implications for *all* medications with improved administration schedules.
  - "Real-world evidence", however, is suggestive of a substantial impact of noncompliance and nonpersistence on treatment effectiveness, notably in ADHD

A Critique (2)

- **Data synthesis across studies and endpoints**
  - Remaining evidence was insufficient to assess relative value of treatment options
  - Synthesis of response rates derived from heterogeneous endpoints (CGI-I / CGI-S vs. narrow-band symptom scales; definitions of response and subscales used)
  - Synthesis of data from heterogeneous studies (including, but not limited to, pooling data from pragmatic "real-world" studies and from double-blind RCTs)
- **Economic model**
  - Not transparent, at times enigmatic description (inclusion of studies, data extracted from studies [e.g., MTA], implausible QALY estimates)
  - Interpreting symptom scales explicitly as "quality of life instruments"
  - Extended time horizon of 12 years without considering long-term sequelae (confounded by technical anomalies, e.g., discount rates applied)
- **Appraisal**
  - The Appraisal Consultation Document noted the ADHD core signs of inattention, hyperactivity, and impulsiveness, the difference between ICD-10 and DSM-IV definitions, and the potential influence of comorbidity on therapeutic outcomes in ADHD, although the Assessment Report had not adequately addressed those[1]

[1]Of note, the appraisal process resulted in a correction of the "clear conclusion" of the assessment report

CASE STUDY: INSIGHTS FROM OTHER HTAs AND CEAs

- **Overview of HTAs and CEAs Related to ADHD**
  - Comparative Evaluations Reporting ICERs
- **NIMH MTA Study**
  - Treatment Modalities
  - Treatment Intensity
- **Compound-Specific Data**
  - Compliance Models
  - Comparative Data

CRITICAL REVIEW - COMPARATIVE EVALUATIONS ONLY

Type	Basis	Agency / Authors	Juris-diction	Comparison	Effectiveness Measure
HTAs	Literature review and decision model	CCOHTA, December 1998 (Zupancic et al., 1998)	CAN	MPH-IR, DEX, PEM; BEH, Comb, NoTx	CTRS (Effect Size / WMD)
		NICE, July 2000 (Lord and Paisley, 2000)	UK	MPH-IR, NoTx	QALYs; (also CTRS points)
		NICE, March 2006 (King et al., 2004, 2006)	UK	DEX, MPH (-IR, -MR08, -MR12), ATX	QALYs based on synth'd. response rates
CEAs	NIMH MTA* Study (1999)	Jensen et al., 2004, 2005	US	CC, BEH, MedMgt, Comb	SNAP-IV Normalization Rates
		Foster et al., 2005, 2006	US	CC, BEH, MedMgt, Comb	Columbia Impairment Scale (CIS)
		Schlander et al., 2004, 2005	US, D	CC, BEH, MedMgt, Comb	SNAP-IV Normalization Rates
	Literature review, model	Narayan and Hay, 2004	US	MPH-IR, MAS[1], NoTx	QALYs based on response rates
	Literature, expert opinion	Iskedijan et al., 2003	CAN	MPH-IR, ATX	SFDs (symptom free days)
	CCOHTA model (ext'd.)	Annemans and Ingham, 2002	CAN	MPH-MR12, MPH-IR (w/ or w/o NDT?)	CPRS (Effect Size)
	Meta-analysis and model	Donnelly et al., 2004	AUS	MPH-IR, DEX	YLD[2]; DALYs (averted)
	Literature review	Gilmore and Milne, 2001 (Wessex DEC Report 1998)	UK	MPH-IR, Plac.	QALYs based on response rates
	Meta-analysis and decision analytic model (CCOHTA ext'd.)	Schlander et al., 2004	UK	MPH-MR12, MPH-IR (w/ NDT)	CTRS (Effect Size)
		Schlander et al., 2004	D	MPH-MR12, MPH-IR (w/ NDT)	CTRS (Effect Size)

[1]MAS: mixed amphetmaine salts

[2]YLD: years lived with disability

The NIMH MTA Study[1]

- **Randomized Clinical Trial of Treatment *Strategies***
  - Psychosocial Treatment Alone [BEH]
  - Pharmacological Treatment Alone [MM]
  - Combined Psychosocial and Pharmacological Treatment [COMB]
  - Community Comparison Group [CC]
- **579 subjects**
  - entered between January and May of three consecutive years
  - six sites (in the United States and Canada)
- **Treatment for 14 months**, follow-up data for +22 months
- **Extensive standardization**
  - Treatment manuals
  - Coordinated staff training
  - Extensive measures of treatment fidelity for all components

[1]MTA Cooperative Group 1999a, 1999b

Effectiveness Data

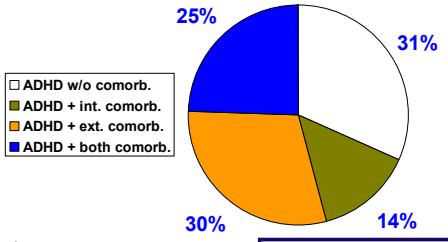
- **Response Rates (SNAP-IV Normalization)**
  - Narrow band symptom scale, integrating parent and teacher scores
  - Capturing DSM-IV defined core symptoms of ADHD (inattention, hyperactivity/impulsivity; also opposition/defiance)[1]
- **Quality-Adjusted Life Year (QALY) Estimates**
  - Response rates defined by symptomatic normalization (SNAP-IV)
  - Health-related quality of life ("utility") weights derived from
    - Expert estimates ("best case" analysis): $\Delta = 0.117$[2]
    - Parent proxy ratings ("base case" analysis): $\Delta = 0.064$[3]
  - Note underlying normative assumption ("extrawelfarism") of QALY maximization; "a QALY is a QALY is a QALY"…
- **Columbia Impairment Scale (CIS) Scores**
  - Global measure of impairment, tapping four domains: interpersonal relations, psychopathology, (job or) schoolwork, use of leisure time

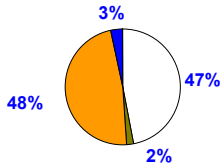
[1]J. Swanson et al. 2001; [2]Lord, S. Paisley 2000; [3]D. Coghill et al. 2004

MTA based economic evaluation of ADHD treatment strategies

Study Population

	ADHD DSM IV								HKD/HKCD ICD10							
Pure ADHD	Total 184								Total 68							
	CC	42	MedMgt	46	Beh	43	Comb	53	CC	13	MedMgt	16	Beh	18	Comb	21
ADHD & Internalizing	Total 81								Total 3							
	CC	19	MedMgt	20	Beh	23	Comb	19	CC	0	MedMgt	0	Beh	3	Comb	0
ADHD & Externalizing	Total 136								Total 69							
	CC	54	MedMgt	40	Beh	42	36		CC	19	MedMgt	17	Beh	19	Comb	14
ADHD & Both Comorbidities	Total 142								Total 5							
	CC	31	MedMgt	38	Beh	36	Comb	37	CC	1	MedMgt	3	Beh	1	Comb	0
Total	Total 579								Total 145							
	CC	145	MedMgt	144	Beh	144	Comb	146	CC	33	MedMgt	36	Beh	41	Comb	35





[1]M. Schlander et al. (2004, 2005)

MTA based economic evaluation of ADHD treatment strategies

Primary Cost-Effectiveness Analysis

Cost per Patient "Normalized" [US-$]

Diagnosis		DSM-IV				ICD-10
Comorbidity	MTA overall	ADHD only	ADHD+intern.	ADHD+extern.	ADHD+both	HKD/HKCD
Comparison						
MedMgt vs. CC	**352**	**dominant**	**869**	**137**	**1,000**	**124**
COMB vs. MedMgt	**55,392**	**48,915**	**inferior**	**75,978**	**29,439**	**31,445**
BEH vs. CC	65,744	47,749	27,245	inferior	22,737	113,462
COMB vs. CC	15,712	14,071	12,062	15,319	13,020	14,350
COMB vs. BEH	2,468	936	4,831	2,090	4,235	2,535
BEH vs. MedMgt	inferior	inferior	inferior	inferior	inferior	inferior

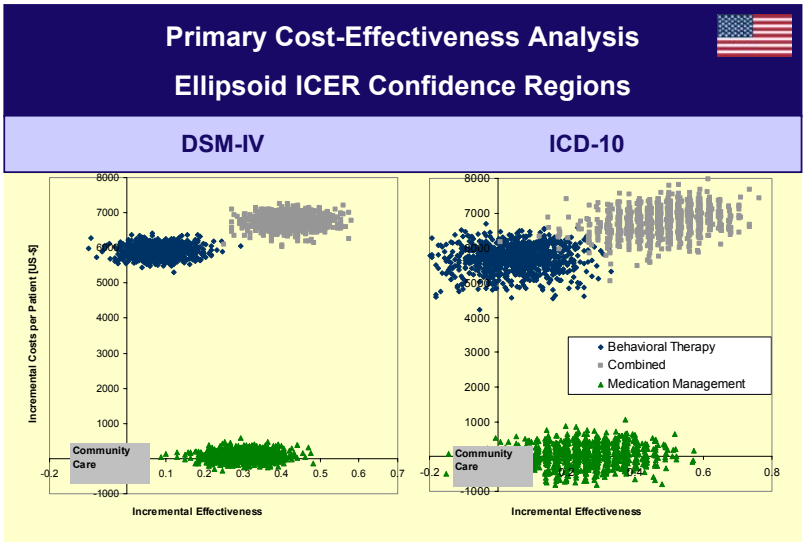
Estimated Cost per QALY Gained [US-$]

(a) Best Case:			**n.a.**	**n.a.**	**n.a.**	
MedMgt vs. CC	**3,009**	**dominant**	**n.a.**	**n.a.**	**n.a.**	**1,060**
COMB vs. MedMgt	**473,436**	**418,077**	**n.a.**	**n.a.**	**n.a.**	**268,761**
BEH vs. CC	561,915	408,111	n.a.	n.a.	n.a.	969,761
COMB vs. BEH	21,094	8,000	n.a.	n.a.	n.a.	21,667
(b) Base Case:						
MedMgt vs. CC	**5,500**	**dominant**	**n.a.**	**n.a.**	**n.a.**	**1,938**
COMB vs. MedMgt	**865,500**	**764,297**	**n.a.**	**n.a.**	**n.a.**	**491,328**
BEH vs. CC	1,027,250	746,078	n.a.	n.a.	n.a.	1,772,844
COMB vs. BEH	38,563	14,625	n.a.	n.a.	n.a.	39,609

[1]M. Schlander et al. (2004, 2005)

CRITICAL REVIEW

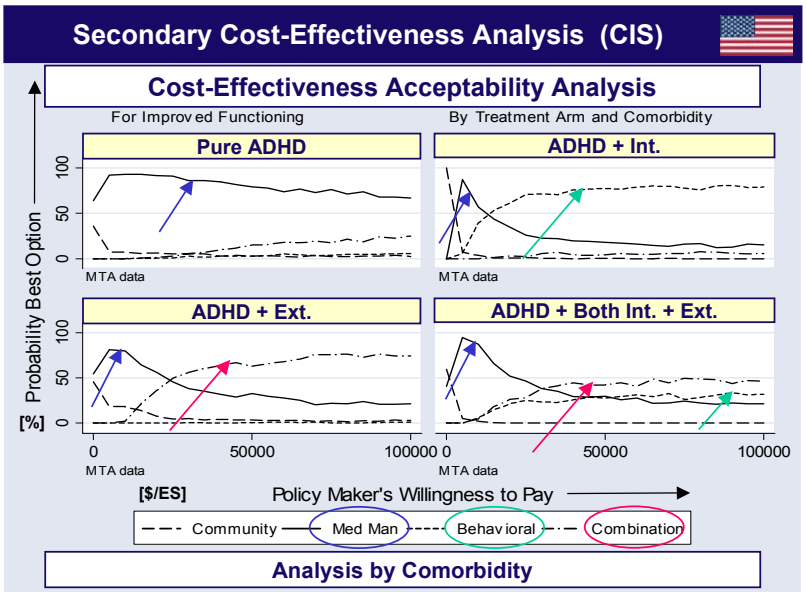
MTA based economic evaluation of ADHD treatment strategies: probabilistic sensitivity analysis



[1]M. Schlander et al. (2005)

CRITICAL REVIEW

MTA based economic evaluation of ADHD treatment strategies



E.M. Foster, P.S. Jensen, M. Schlander, et al. (2005)

Conclusions based upon the NIMH MTA Study

- An intense medication management strategy appears attractive on grounds of cost-effectiveness[1].
- This conclusion holds for the subgroup of patients meeting ICD-10 criteria for hyperkinetic disorder.
- Ambitious psychosocial interventions are unlikely to meet benchmarks for cost-effectiveness.
- There is a strong research need to determine the role of better tailored psychosocial interventions.
- Comorbidity is an important moderator of cost-effectiveness of ADHD treatment strategies, contingent on therapeutic objectives.

[1]Note that this data provide economic insights about cost-effectiveness "at the margin"

Comparing Formulations (Addressing Compliance)[1]

- **Real world studies**
  - Although not without limtations itself, the randomized open-label real-world study included in the economic model reported greater effect differences in favor of modified-release MPH than found after pooling *all* studies.
- **Decision analytic models**
  - Replicates (CAN,UK, D) of the original CCOHTA model (with symptomatic improvements on the Conners scales as effectiveness criterion and a range of compliance assumptions) indicate acceptable cost-effectiveness (and the possibility of extended dominance) of MPH-MR[2] compared to MPH-IR t.i.d.
- **Empirical evidence**
  - Empirical evidence has now become available lending support to some of the assumptions of the modeling studies. Three claims database analyses in the U.S. showed that ADHD treatment persistence with MPH-MR[2] exceeded that achieved with immediate-release formulations. Data from one analysis further indicate that this observation cannot be explained simply by patient selection bias.

[1]multiple references
[2]data currently limited to MPH-MR12

Comparing Compounds[1]

- **Head-to-head studies**
  - Evidence consists of two direct comparisons of atomoxetine and MPH-MR[2]. One of these trials was erroneously missed in the assessment.
  - Both trials showed superior effectiveness of MPH-MR[2], but there existed issues related to study inclusion criteria (resulting in inclusion of few stimulant-naive patients, likely creating a bias in favor of MPH-MR). In subgroup analyses looking at stimulant-naive patients only, the difference was still present but no longer statistically significant.
- **Effect size differences**
  - In an analysis of effect sizes achieved with once-daily administration based on Conners scores (indicating symptomatic improvement), MPH-MR[2] (ES ~1.0) was superior to atomoxetine (ES ~0.7); this analysis was not considered for assessment.
- **Other technology assessments**
  - Assessing the same data as NICE, the Scottish Medicines Consortium (SMC) did not recommend atomoxetine, because "the economic case has not been demonstrated". PBAC reached a similar conclusion, refusing PBS listing of atomoxetine.

[1]multiple references
[2]data currently limited to MPH-MR12

Available Clinical Evidence Not Fully Used

- **Data using Conners ratings**
  - The most widely used group of scales in ADHD studies, with well-established psychometric properties – would have enabled access to "long-term" data.
  - Extensive literature on instruments to measure clinical outcomes in ADHD
- **Compliance models**
  - Including data on noncompliance in ADHD
  - Extensive compliance research literature
- **Real-world effectiveness data**
  - Pooled with efficacy data from RCTs
- **Head-to-head comparisons**
  - Incomplete search for evidence

HAS NICE GOT IT RIGHT?

- Praise for Approach
- Praise for Process
- Research Question
- Motivation and Limitations
- Qualitative Case Study: TA No. 98
- **Accountability for Reasonableness**
- **Some Suggested Underlying Issues**
- **Some Implications**

HAS NICE GOT IT RIGHT?

Technology Appraisal of Methylphenidate, Dexamphetamine and Atomoxetine (NICE 2006)

Accountability for Reasonableness[1]

- **Publicity:**
  - Decisions and their underlying rationales must be publicly accessible.
- **Relevance:**
  - These rationales must rest on evidence, reasons, and principles that plan managers, clinicians, patients, and consumers agree are pertinent to deciding how to meet diverse needs under resource restraints.
- **Revisability and appeals:**
  - A mechanism must allow challenges to limit-setting decisions, help resolve those challenges, and allow revisions in light of further evidence and arguments.
- **Enforcement:**
  - A voluntary or public regulatory process must ensure that decision makers fulfill the first three conditions.

[1]Source: N. Daniels, J.Sabin (1998ff.)

Symptoms and some suggested underlying issues

Separation of Clinical and Economic Evaluation

Differences in scope

Selection of clinical studies
(interpretation of 3-weeks-duration criterion;
absence of consideration of carry-over effects in crossover trials)

Dissociation between effectiveness review and cost-effectiveness evaluation of technology assessment,
the latter not using findings of the systematic review
(i.e., use of hyperactivity scores versus CGI-I subscale scores)

Disorder-specific outcome measures
not considered for economic evaluation,
contributing to the exclusion from analysis of **clinical long-term evidence**
(including absence of literature review on clinical measurement instruments and on long-term outcomes)

Broad use of secondary endpoints of clinical studies
as an input for probabilistic cost-effectiveness evaluations
(intended to capture stochastic uncertainty)

Distinction between efficacy and effectiveness
(including absence of compliance literature review)

Reasoning that utility values obtained directly from patients
"may be more relevant to this review",
raising doubt whether the clinical problem was fully understood by analysts

Symptoms and some suggested underlying issues

High Level of Standardization

Exclusive focus on cost-utility analyses

- At the expense of cost-effectiveness evaluations
- Reliance on utility estimates of limited validity
- For calculation of quality-adjusted life years (QALYs),
  linking utility estimates based on complex health state descriptions
  with response estimates based on clinical global impressions subscales
- Inability to identify differences between treatments

Highly restrictive use of clinical evidence for economic evaluation

- Clinical long-term studies
- Commonly used effectiveness measures
- Mathematical precision of quantitative meta-analysis
  not in tune with imprecision of dichotomized input data
  (mostly CGI-based "response rates",
  or data pooled from heterogeneous sources)
  from small-scale short-term clinical studies
- Need to use data from clinical studies
  that had been excluded from effectiveness review for quality concerns

Symptoms and some suggested underlying issues

Need for (or Absence of Effective) Quality Assurance

Deviation of assessment from NICE reference case

- Discount rates used for long-term economic model
- Discussion of appropriate sources of utility estimates

Issues related to technical quality of assessment

- Multiple deviations from specified search criteria
  (relevant randomized clinical studies;
  relevant health economic evaluations;
  interpretation of 3-weeks cut-off criterion;
  inclusion of studies previously rejected for quality concerns)
- Pooling of heterogeneous studies for quantitative synthesis
  (e.g., efficacy vs. effectiveness; clinical effectiveness measures,
  treatment intensity, concomitant psychosocial treatment, etc.)
- Not controlling for potential confounding effects
  (e.g., effectiveness measures used and treatment strategies)
- Mismatch between clinical global impressions
  (and other response criteria used)
  and health state descriptions used for utility estimates

Technology Appraisal of Methylphenidate, Dexamphetamine and Atomoxetine (NICE 2006)

Has NICE Got It Right *Consistently*?

- **Apparently, the answer is "Not really".**
  - The current NICE approach to health technology appraisals, although often considered exemplary from an international perspective, may become overstretched by complex clinical problems.
- **Suggested underlying reasons include:**
  - Insufficient integration of clinical and economic evaluation.
  - High level of standardization, contributing to a relatively rigid application of the cost-utility (cost-per-QALY) concept, at the expense of alternative methods of health economic evaluation.
  - Provisions for (or lack of) quality assurance for technology assessments.
  - Some process-related issues (primarily related to the relevance condition of A4R and the use of "QALY egalitarianism" as fundamental equity position, contributing to NICE's strong focus on QALYs).

Does this observation invalidate the approach taken by NICE?

- **Again, the suggested answer is "No" ("Not really").**
  - One qualitative case study (n=1) does not allow inferences about N>100 technology appraisals by NICE.
  - Of note, the NICE appraisal **process** enabled correction of some of the observed limitations of the technology assessment.
  - Nevertheless, qualitative research exploring specific issues *in-depth* may create hypotheses that deserve further research.
  - There are indeed some indications that certain problems observed in the present case might not have been a singular occurrence.
  - Given the impact of NICE guidance, the limitations associated with the assessment of ADHD treatment strategies are considered serious enough to warrant further inquiry.

An International Perspective

- **International policy makers, looking at NICE as a potential role model, might wish to consider...**
  - **... Objectives** of health care provision (and collective financing); which criteria are appropriate to determine "allocative efficiency" in line with social values?
  - **... Institutional context** of the respective health care system.
  - **... Reliance on QALYs** as an appropriate outcome measure?
  - **... Which technology appraisal processes**? (With few exceptions, it is suggested here that NICE might indeed serve as a role model in that respect.)
  - **... Timing** of technology appraisals?
  - **... Multidisciplinary assessment teams**?
  - **... Quality assurance** of technology assessments?
  - **... Implementation issues**...