Psychological Medicine

cambridge.org/psm

Long-term effects of stimulant treatment on ADHD symptoms, social–emotional functioning, and cognition

Lizanne Schweren[1], Pieter Hoekstra[1], Marloes van Lieshout[2], Jaap Oosterlaan[2], Nanda Lambregts-Rommelse[3], Jan Buitelaar[3,4], Barbara Franke[3,5] and Catharina Hartman[1]

[1]Department of Psychiatry, University Medical Center Groningen, Groningen, The Netherlands; [2]VU University Amsterdam, Amsterdam, The Netherlands; [3]Radboud University Medical Center, Donders Institute for Brain, Cognition and Behaviour, Nijmegen, The Netherlands; [4]Karakter Child and Adolescent Psychiatry, Nijmegen, The Netherlands and [5]Department of Human Genetics, Radboud University Medical Center, Nijmegen, The Netherlands

Original Article

Cite this article: Schweren L, Hoekstra P, van Lieshout M, Oosterlaan J, Lambregts-Rommelse N, Buitelaar J, Franke B, Hartman C (2018). Long-term effects of stimulant treatment on ADHD symptoms, social–emotional functioning, and cognition. *Psychological Medicine* **49**, 217–223. https://doi.org/10.1017/S0033291718000545

Received: 18 January 2017
Revised: 12 December 2017
Accepted: 12 February 2018
First published online: 13 March 2018

Key words:
Attention; attention-deficit/hyperactivity disorder; cognition; methylphenidate; pharmacotherapy; stimulant treatment

Author for correspondence:
Lizanne Schweren, E-mail: l.j.s.schweren@umcg.nl

Abstract

Background. Methodological and ethical constraints have hampered studies into long-term lasting outcomes of stimulant treatment in individuals with attention-deficit/hyperactivity disorder (ADHD). Lasting effects may be beneficial (i.e. improved functioning even when treatment is temporarily ceased) or detrimental (i.e. worse functioning while off medication), but both hypotheses currently lack empirical support. Here we investigate whether stimulant treatment history predicts long-term development of ADHD symptoms, social–emotional functioning or cognition, measured after medication wash-out.

Methods. ADHD symptoms, social–emotional functioning and cognitive test performance were measured twice, 6 years apart, in two ADHD groups (stimulant-treated *versus* not stimulant-treated between baseline and follow-up). Groups were closely matched on baseline clinical and demographic variables ($n = 148$, 58% male, age = 11.1). A matched healthy control group was included for reference.

Results. All but two outcome measures (emotional problems and prosocial behaviour) improved between baseline and follow-up. Improvement over time in the stimulant-treated group did not differ from improvement in the not stimulant-treated group on any outcome measure.

Conclusions. Stimulant treatment is not associated with the long-term developmental course of ADHD symptoms, social–emotional functioning, motor control, timing or verbal working memory. Adolescence is characterised by clinical improvement regardless of stimulant treatment during that time. These findings are an important source to inform the scientific and public debate.

Introduction

Attention-deficit/hyperactivity disorder (ADHD) is a prevalent and often persistent developmental disorder, characterised by age-inappropriate and impairing levels of inattention and/or hyperactivity–impulsivity. ADHD has been associated with a broad range of neurocognitive deficits, including impaired executive functioning (Willcutt *et al.* 2005), timing deficits (Noreika *et al.* 2013) and higher response time variability (Klein *et al.* 2006). In the majority of individuals with ADHD, stimulants acutely reduce symptoms (Swanson *et al.* 2001) and improve neurocognitive functioning (Coghill *et al.* 2014). Concerns about potential harmful long-term effects of stimulant treatment, as well as anticipation of potential lasting benefits of treatment have dominated the public and scientific debate. Adequately investigating long-term treatment effects, especially in children, is methodologically and ethically challenging, hence evidence for either positive or negative long-term outcomes of stimulant treatment is equivocal. In the Multimodal Treatment Study of ADHD (MTA), the largest controlled treatment study to date, the benefits of 14 months of stimulant treatment on a broad range of outcomes rapidly diminished in the subsequent observational phase (MTA Cooperative Group, 1999; Swanson *et al.* 2007; Molina *et al.* 2009). In the MTA study, outcomes were assessed without a medication wash-out phase, which impedes the distinction between lasting effects of prior treatment and acute effects of ongoing treatment. When rated while off-medication, ADHD symptoms were found not to change with 1 year of stimulant treatment (Huang *et al.* 2012). Attention task performance and IQ did improve over the course of 1 year, but in the absence of a comparable non-treated or healthy control group, these changes may reflect normal maturation (Tsai *et al.* 2013). Observational studies have reported higher ADHD

CAMBRIDGE UNIVERSITY PRESS

persistence rates in stimulant-treated patients compared with non-treated patients (Biederman *et al.* 2012; van Lieshout *et al.* 2016), while at the same time, rates of comorbidity were found to be lower in treated patients (Biederman *et al.* 2009). Importantly, in these studies confounding-by-indication and self-selection could not satisfactorily be addressed. Here, we applied stringent matching procedures to derive two comparable ADHD samples from a large prospective cohort study (i.e. stimulant-treated and not stimulant-treated) as well as a typically developing reference group. Outcomes were repeatedly measured over 6 years, always while participants were in their non-medicated state. We investigated whether stimulant treatment between baseline and follow-up predicted the developmental trajectory of ADHD symptoms, social–emotional functioning and/or cognitive functioning in the following domains: motor control, timing and verbal working memory. Note that, while there is evidence that these domains are affected in ADHD and may benefit (acutely) from methylphenidate treatment (Rubia *et al.* 2003; Yang *et al.* 2012; Kaiser *et al.* 2014), our choice of cognitive domains was limited by task availability. Tasks measuring response inhibition, reward sensitivity and visuospatial (rather than verbal) working memory would have been highly informative but unfortunately were unavailable within the current sample.

Methods

Participants

Participants were drawn from the prospective multi-centre *IMAGE-NeuroIMAGE* cohort study (von Rhein *et al.* 2015). The full cohort includes 751 children, adolescents and young adults with ADHD from 590 families. At baseline, ADHD diagnosis was ascertained using the Strengths and Difficulties Questionnaire (SDQ; >90th percentile on the hyperactivity subscale; van Widenfelt *et al.* 2003), the parent- and teacher-rated Conners' ADHD scales (CPRS and CTRS; $T \geqslant 63$ on the DSM inattentive or hyperactive/impulsive scale; Conners *et al.* 1998*a*, *b*) and the Parental Account of Children's Symptoms interview (PACS; ⩾6 symptoms, present in ⩾2 situations and ⩾1 symptom reported by the teacher; Tailor, 1986). Participants with ⩾6 symptoms but who did not fulfil all diagnostic criteria were classified as subthreshold ADHD. At follow-up, ADHD diagnosis in participants <18 years was ascertained again using the same CPRS and CTRS criteria, complemented with the Schedule for Affective Disorders and Schizophrenia for School-Age Children interview (K-SADS; ⩾6 symptoms, present in ⩾2 situations, causing impairment, and onset before age 12; Kaufman *et al.* 1997). For participants ⩾18 years, the self-rated Conners' scale (CAARS; Conners *et al.* 1999) was used instead of the teacher-rated scale, and five symptoms were sufficient for diagnosis. Participants who scored $T \geqslant 63$ on either of the Conners' scales or had sufficient symptoms, but did not fulfil all diagnostic criteria, were classified as subthreshold ADHD.

Average follow-up time was 5.9 years (S.D. = 0.6), and the retention rate was high (77%). We applied the following inclusion criteria: (1) participation at baseline and follow-up, (2) diagnosis of (subthreshold) ADHD at baseline and/or at follow-up, (3) IQ>70 at baseline and follow-up, and (4) no known genetic or neurological disorders. Eligible participants were split according to treatment between baseline and follow-up into stimulant-treated ($n = 337$) and not stimulant-treated participants ($n = 138$). Stimulant treatment prior to baseline and treatment with non-stimulant psychoactive medication was allowed in both groups. From the two ADHD groups, we selected all participants who had a one-to-one match on gender, age (±<0.5 S.D.) and baseline number of ADHD symptoms (± < 0.5 S.D.). This resulted in two comparable groups of 74 participants with ADHD each (Table 1).

For reference, a gender- and age-matched healthy control sample was drawn from the *IMAGE-NeuroIMAGE* cohort as well, applying the same inclusion and matching criteria (except inclusion criterion 2/symptom-matching). In addition, control participants had no first-degree relatives with psychiatric disorders, as ascertained by interview. All assessments took place at two sites in the Netherlands. Participants were asked to withhold use of psychoactive drugs for 48 h before each assessment. Informed consent was signed by all participants and their parents (only parents signed informed consent for participants <12 years). Procedures were approved by the local ethical committee of each site.

Stimulant treatment

Participants and parents provided written consent to request prescription records from their pharmacies. In addition, they reported lifetime history of psychoactive medication in a questionnaire at follow-up measurement. Pharmacy data covering the baseline–follow-up interval were available for 91% of participants with ADHD ($n = 135$). Participants were classified as stimulant-treated if they had been prescribed any immediate or extended release methylphenidate preparations, or *d*-amphetamine preparations, between baseline and follow-up. When pharmacy transcripts were not available or incomplete ($n = 13$), treatment history was derived from the questionnaire data. The questionnaire data were also used to determine stimulant treatment prior to baseline ('previously treated' or 'stimulant-naïve') for all participants.

Outcome measures

Parent-rated numbers of hyperactivity–impulsivity and inattention symptoms were measured at baseline and follow-up using the respective DSM subscales of the CPRS (range 0–27). For participants using medication, parents were instructed to rate behaviour in the participant's non-medicated state. Four indicators of social–emotional functioning were derived from the SDQ for both time points: problems with emotion regulation, problems with peer relationships, conduct problems and prosocial behaviour (range 0–10).

In addition, six cognitive tests were administered at both baseline and follow-up. Three tasks measured motor control: baseline speed, in which participants were required to press a key upon unpredictable appearance of a stimulus; pursuit, where participants followed a randomly moving target with the cursor as precisely as possible; and tracking, in which participants were required to trace an invisible midline between an inner and an outer circle as precisely as possible. Two tasks measured timing: time estimation, where participants were asked to reproduce the duration of visually presented stimuli of different lengths (4, 8, 12, 16 and 20 s); and motor timing, in which participants were instructed to produce 1 s intervals as accurately as possible. Working memory was assessed in the backwards condition of the digit span test (WISC-III/WAIS-III), in which participants had to reproduce an increasingly long sequence of numbers in reverse order. Details are in Table 2. We note that several cognitive domains that are relevant to ADHD, including inhibition and delay aversion, were not available at both time points and could not be evaluated here.

https://doi.org/10.1017/S0033291718000545 Published online by Cambridge University Press

Table 1. Baseline characteristics of the two treatment groups

	Treated		Non-treated			
	Mean	S.D.	Mean	S.D.	Stat.	*p*
Gender = male	*N* = 43	58.1%	*N* = 43	58.1%	0.000	1.000
Age	11.14	3.29	11.00	3.23	0.066	0.798
Site = Amsterdam	*N* = 27	36.5%	*N* = 46	62.2%	9.759	0.002[a]
IQ	99.93	10.47	103.55	10.77	3.605	0.060
Socio-economic status	11.26	2.02	12.07	2.52	4.522	0.035[a]
Follow-up interval (years)	5.92	0.60	5.86	0.68	0.258	0.613
Treatment prior to baseline = yes	*N* = 52	70.3%	*N* = 18	24.3%	31.335	<0.001[a]
ADHD type					8.677	0.070
Unaffected	*N* = 6	8.1%	*N* = 7	9.5%		
Inattentive	*N* = 4	5.4%	*N* = 6	8.1%		
Hyperactive	*N* = 1	1.4%	*N* = 2	2.7%		
Combined	*N* = 55	74.3%	*N* = 39	52.7%		
Subthreshold	*N* = 8	10.8%	*N* = 20	27.0%		
Co-morbid problems[b]						
Anxiety/shyness	5.20	4.92	4.30	4.47	1.333	0.250
Perfectionism	3.85	4.24	3.55	3.55	0.214	0.644
Psychosomatic problems	3.45	3.33	2.80	3.16	1.445	0.231

Stat = χ^2 for categorical variables, Student's *t* test for continuous variables.
[a]Significant difference between treated and non-treated participants
[b]Scores on the anxiety/shyness scale, perfectionism scale and psychosomatic problems scale of the parent- and teacher-rated Conners' questionnaires were used as a proxy of baseline comorbid problems.

Statistical analyses

We used linear mixed-effects models, predicting symptoms of hyperactivity/impulsivity and inattention, each of the four social–emotional outcomes, and performance on each cognitive test from time (baseline or follow-up), treatment (stimulant-treated or not stimulant-treated during the study phase) and time-by-treatment interaction. The effect of interest is captured in the time-by-treatment interaction, which evaluated whether the outcome variables changed differently over time for the stimulant-treated group compared with

Table 2. Neurocognitive tasks

Task (aim)	Description	Performance measure	*N*
Baseline speed (motor output in response to cue)	Participants were required to press a key after a white square appeared unpredictably (500–2500 ms after response) on a screen	Standard deviation of reaction times in ms averaged across left and right hand	78 (52.7%)
Pursuit (motor control with continuous adaptation)	Participants were required to 'catch' a randomly moving stimulus (asterisk, 10 mm/s) as precisely as possible by moving the cursor on top of the stimulus with the left hand	Mean absolute distance in mm between target and cursor	81 (54.7%)
Tracking (motor control without continuous adaptation)	With the left hand, participants traced an invisible midline between an inner and outer circle presented on the screen (radius 7.5 and 8.5 cm, respectively), counterclockwise and as quickly and precisely as possible	Mean absolute distance in mm between target (midline) and cursor	83 (56.1%)
Digit span (working memory)	Participants were instructed to reproduce sequences of numbers, of increasing length, in reverse order	Maximum accurately reproduced sequence length	111 (75.0%)
Time estimation	Stimuli (4, 8, 12, 16, 20 s) were randomly presented by a lightbulb. Participants were required to reproduce stimulus length by pressing a button	Absolute discrepancy between the response length and the stimulus length averaged across all 12 s trials	83 (56.1%)
Motor timing	Participants were instructed to produce a 1 s interval after a tone, as accurately as possible. Visual feedback was given (correct, too short or too long; defined by a dynamic tracking algorithm)	Median absolute deviation in ms from 1 s	88 (59.5%)

N = number of participants with ADHD who completed the task at baseline and at follow-up.

https://doi.org/10.1017/S0033291718000545 Published online by Cambridge University Press

the non-treated group. Baseline demographic/clinical between-group differences that remained despite matching [testing site, socio-economic status (SES) and treatment prior to baseline] were included as covariates, as was a random intercept per family to account for dependencies among siblings. Multiple testing was accounted for by Bonferroni adjustment: α was divided by two for ADHD symptoms ($\alpha = 0.05/2 = 0.025$), by four for social–emotional outcomes ($\alpha = 0.012$) and by six for cognitive outcomes ($\alpha = 0.008$).

Previous work by our group described changes over time in ADHD symptoms in participants with ADHD compared with typically developing participants (van Lieshout *et al.* 2016). Case–control differences are thus not the focus of the current study. Rather, the matched control group was used as a reference group for normative developmental changes. For visualisation of estimated marginal means of all groups (stimulant-treated, not stimulant-treated and control), the models described above were re-estimated across all participants with a fixed factor for group.

Sensitivity analyses were performed to test the robustness of our findings. With a relatively short wash-out time (48 h), immediate withdrawal effects may have affected cognitive functioning in participants who received ongoing treatment at the time of measurement. Therefore, analyses were repeated with an additional covariate encoding whether participants were actively being treated with stimulants within 6 months prior to assessment or not, and its interaction with the effect of interest (active treatment × time × treatment between baseline and follow-up). Second, all analyses were repeated with baseline age as an additional predictor, to address the wide age range within our sample. Here, similarly, change over time in each outcome variable was predicted from age-by-treatment interaction, thus analysing whether the effect of treatment on clinical/social–emotional/cognitive changes over time was different for participants of different ages.

Results

Mean age of participants with ADHD was 11.1 years (S.D. = 3.2) at baseline and 17.0 years (S.D. = 3.3) at follow-up. Fifty-eight per cent of participants were male. Participants were diagnosed with ADHD or subthreshold ADHD at baseline ($n = 135$, 91.2%) and/or at follow-up ($n = 132$, 89.2%). Most participants reached diagnostic criteria at both times ($n = 119$, 80.41%). Fifteen participants (10.1%) with subthreshold ADHD never met criteria for full ADHD diagnosis. At baseline, the majority of participants had combined type ADHD ($n = 94$, 63.5%), while at follow-up, the majority had either combined type ($n = 40$, 27.0%) or inattentive type ($n = 51$, 34.5%), with no differences between groups (Table 1). Within the stimulant-treated group, average cumulative stimulant dose between baseline and follow-up was 43 336 mg, which equals 5.9 years of 20.1 mg per day. Forty participants (54.1%) had received active stimulant treatment within 6 months prior to follow-up assessment; the other participants had ceased stimulant treatment earlier. Participants in the stimulant-treated group were from lower socio-economic backgrounds ($p = 0.035$), were more likely to have received stimulant treatment prior to the initial assessment ($\chi^2 = 31.335$, $p = 0.001$) and more likely to have received atomoxetine treatment between baseline and follow-up ($n_{OVERALL} = 16$, 10.8%; $n_{TREATED} = 13$, 17.6%; $n_{NON-TREATED} = 3$, 4.1%; $\chi^2 = 6.862$, $p = 0.009$). There was a site effect for stimulant treatment as well ($\chi^2 = 9.759$, $p = 0.002$). Site, SES and prior treatment were therefore added as covariates in all between-group comparisons. At baseline, the two treatment groups did not differ from each other with regard to any of the clinical or cognitive outcome measures.

There was a significant main effect of time on ADHD symptoms, as well as on two out of four social–emotional outcome measures (Table 3). Across all participants with ADHD, symptoms of hyperactivity/impulsivity and inattention, peer problems and conduct problems improved between baseline and follow-up. There were no main effects of time on emotional problems or prosocial behaviour. Improvement over time was also found for performance on all cognitive tasks: participants showed lower baseline speed variability, smaller deviations on the tracking, pursuit and time estimation tasks, and higher maximum digit span at follow-up compared with baseline. Potential confounder's site and

Table 3. Baseline and follow-up scores across treatment groups, and the effects of time, treatment and time-by-treatment interaction

	Baseline		Follow-up				
	EMM	S.D.	EMM	S.D.	p_{TIME}	$p_{TREATMENT}$	$p_{TIME\times TREATMENT}$
Hyperactivity/impulsivity symptoms	14.22	5.95	11.83	6.73	<0.001*	0.212	0.188
Inattention symptoms	12.28	6.15	7.38	5.55	<0.001*	0.557	0.054
Emotional problems	2.98	3.00	2.82	3.08	0.736	0.577	0.707
Prosocial behaviour	7.15	2.08	7.38	2.19	0.351	0.280	0.142
Peer problems	2.82	2.12	2.19	1.98	0.003*	0.382	0.424
Conduct problems	3.09	2.00	2.43	1.83	0.002*	0.238	0.906
Baseline speed variability	172.37	103.89	90.29	50.35	<0.001*	0.513	0.672
Pursuit (inaccuracy)	6.44	3.74	3.87	0.76	<0.001*	0.609	0.320
Tracking (inaccuracy)	2.85	1.81	1.34	0.94	<0.001*	0.798	0.175
Motor timing (inaccuracy)	203.11	95.10	148.83	51.48	<0.001*	0.449	0.341
Time estimation (inaccuracy)	2.72	1.79	1.48	0.81	<0.001*	0.776	0.411
Digit span	3.92	1.15	4.49	1.26	<0.001*	0.126	0.715

EMM = estimated mean score across participants with ADHD, adjusted for stimulant treatment prior to baseline measurement, site and SES.
*$p < 0.012$ or $p < 0.008$.

https://doi.org/10.1017/S0033291718000545 Published online by Cambridge University Press

SES had no main effect on any outcome with one exception: lower SES was nominally associated with more peer problems ($t = -2.340$, $p = 0.021$).

There were no main effects of treatment group, and no time-by-treatment-group interaction effects on any of the outcome measures (Table 3, Fig. 1 and Fig. 2). Thus, changes in ADHD symptoms, social–emotional and cognitive functioning over time were the same for participants who received stimulant treatment between baseline and follow-up and those who had not. Moreover, changes over time were the same for participants on active stimulant treatment at follow-up assessment and those who were not, suggesting no confounding by withdrawal effects. Finally, there were no significant interactions with age, site or SES, suggesting that treatment effects were similar at different levels of these covariates.

Discussion

We investigated developmental changes in a broad spectrum of outcomes, including ADHD symptoms, social–emotional functioning and cognition, in two groups of individuals with ADHD defined by whether they had been treated with stimulants or not. The groups were stringently matched on baseline characteristics and were non-medicated at both assessments. We found no evidence for any (beneficial or adverse) stimulant treatment effects persisting after stimulant treatment had (temporarily) been ceased. ADHD symptoms, peer problems, conduct problems and performance on tests of motor control, timing and working memory improved over time, but improvement occurred irrespective of treatment. Even at a lenient threshold for statistical significance, stimulant treatment was not associated with any of the outcomes.

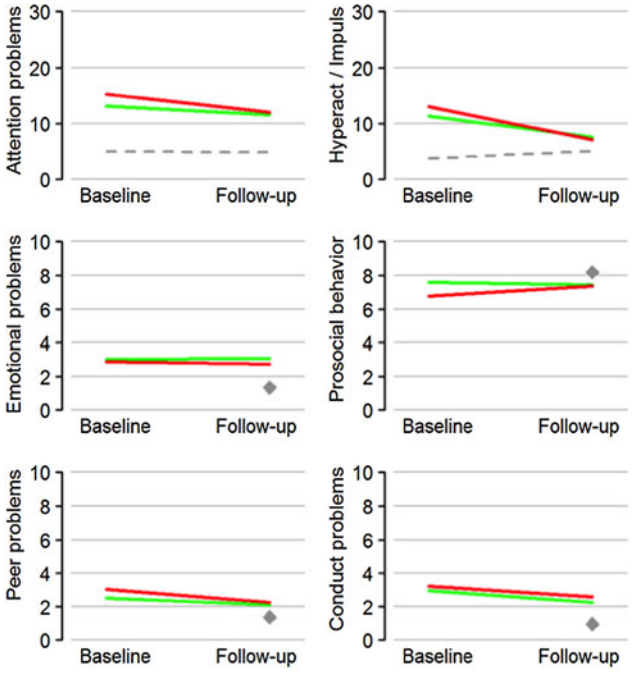

Fig. 1. Change in ADHD symptoms and social–emotional functioning over ~6 years, for stimulant-treated (green) and non-treated (red) participants with ADHD, and control participants (grey). Groups are matched on baseline age and gender, and ADHD groups are matched on baseline ADHD symptoms. Baseline social–emotional outcomes were not assessed for typically developing participants. The slopes of the two treatment groups did not differ for any outcome.

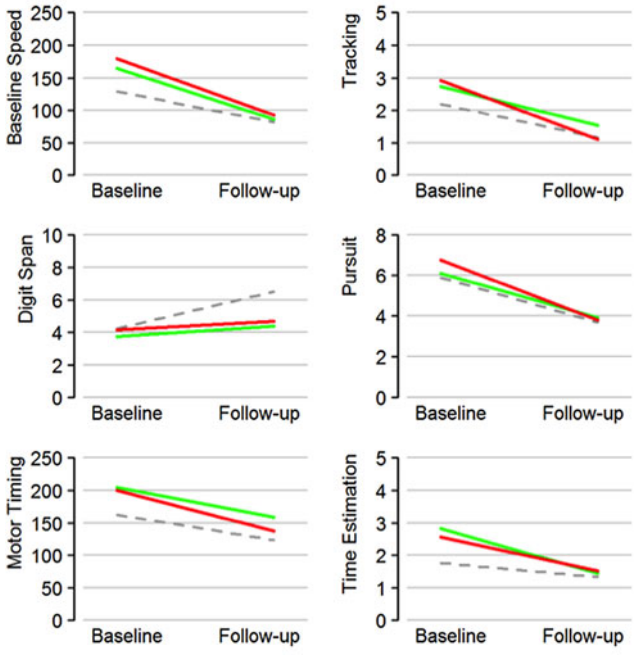

Fig. 2. Change in cognitive test performance over ~6 years, for stimulant-treated (green) and non-treated (red) participants with ADHD, and control participants (grey). Groups are matched on baseline age and gender, and ADHD groups are matched on baseline ADHD symptoms. The slopes of the two treatment groups did not differ for any outcome.

Our findings put into perspective previous studies reporting potential beneficial long-term effects of stimulant treatment that did not allow firm conclusions. First, previous studies reporting long-term beneficial treatment effects typically assessed outcomes when patients were on active treatment (e.g. Abikoff *et al.* 2004; Charach *et al.* 2004). Their findings may thus represent either lasting effects of prior treatment, transient effects of ongoing treatment or a combination of both. Our findings, in conjunction with reports of better outcome during phases of active stimulant treatment (Lichtenstein *et al.* 2012; Chang *et al.* 2016) suggest that previously reported long-term effects may be driven by ongoing transient effects rather than lasting effects. The absence of lasting treatment effects in our sample aligns with negative long-term findings during the observational phase of the MTA study, which have previously been attributed to self-selection (Swanson *et al.* 2007; Molina *et al.* 2009). Our findings, however, underline the possibility that the theorised long-term effects may in fact not occur. At the same time, we wish to emphasise that beneficial long-term treatment effects have been found in outcomes that were not addressed here, such as the development of comorbid disorders later in life (Biederman *et al.* 2009).

Second, our findings are in line with a previous report of improved attention task performance after a 1-year stimulant treatment episode even while off medication (Huang *et al.* 2012), which, in the absence of a reference group, could indicate either lasting beneficial treatment effects or improved cognitive performance at older age. In the current study, changes over time were the same in the treated and non-treated groups, suggesting that improvement over time is not related to treatment.

Third, several previous studies have reported more severe and/or more persistent ADHD in individuals who had received stimulant treatment during childhood, which could indicate either

https://doi.org/10.1017/S0033291718000545 Published online by Cambridge University Press

detrimental treatment effects or confounding-by-indication (Molina *et al.* 2009; Biederman *et al.* 2012; van Lieshout *et al.* 2016). The current findings, free of confounding-by-indication due to stringent matching procedures and accounting for baseline measurements, provide no evidence of detrimental treatment effects.

The current findings are an important source to inform the scientific and public debate about pharmacological treatment for ADHD that has focused on long-term hazards and benefits. First, our findings emphasise that the developmental course of ADHD symptoms, social–emotional outcomes and cognitive functioning, at least for the areas of cognition assessed in the current study, are not altered by stimulant treatment. Previous work of our group showed that ADHD symptoms tend to decline but not disappear at later age (van Lieshout *et al.* 2016). The current results add to these findings by showing that this conclusion holds for both stimulant-treated and non-treated individuals. Second, the absence of long-term treatment effects on clinical and selected cognitive outcomes may guide the interpretation of findings of structural brain changes associated with stimulant treatment (or the absence thereof). The evidence for such an association is mixed (Shaw *et al.* 2009, 2014; Schweren *et al.* 2015). The absence of lasting treatment effects on a broad spectrum of clinical/behavioural outcomes emphasises the importance of investigating behavioural correlates and clinical relevance of stimulant effects on the brain.

This is the first longitudinal study investigating long-term treatment effects that included a non-treated ADHD and a typically developing sample, and reported on a wide spectrum of clinical and cognitive outcomes while participants were non-medicated. The average follow-up time of almost 6 years allowed the detection of effects emerging at later age, and captured the late adolescent/early adulthood phase that is often characterised by both clinical and normative developmental changes, which we were able to tease apart. Our rigorous one-to-one matching procedure allowed firm conclusions. Finally, extensive diagnostic assessments resulted in a well-characterised ADHD sample, and the availability of pharmacy records enabled highly reliable assessment of treatment history.

The current study had limitations as well. Treatment allocation was not random. We were able to rule out confounding-by-indication for all *measured* baseline variables other than testing site and SES, but not for *non-measured* potential between-group differences. Especially functional impairment and comorbidity could not satisfactorily be addressed. Propensity score adjustment would have been valuable in this regard. Confounding may also have occurred during the study phase, e.g. behavioural treatment (not assessed) may have been more common in one group compared with the other. Second, inclusion of matched participants was based on the smallest ADHD group, i.e. those who did not receive stimulant treatment between baseline and follow-up. This may have resulted in a sample that is less representative of the ADHD population. For example, the number of symptoms in the current sample was slightly lower, and the number of females was slightly higher, compared with the full sample as described elsewhere (von Rhein *et al.* 2015); the rate of symptom change between baseline and follow-up, however, was the same (data not shown). Third, cognitive domains that are pertinent to ADHD were not consistently assessed across time. Long-term changes may have occurred in these domains, while long-term changes may have been less likely to occur within the domains we were able to evaluate (e.g. verbal rather than visuospatial working memory). Finally, the current design did not allow full investigation of treatment timing, since participants had often initiated treatment prior to the baseline measurement and/or continued treatment after the follow-up measurement. Treatment at different ages may be associated with different long-term consequences, although in our sample, we found no indications of such effects.

In conclusion, we find no evidence that stimulant treatment may have a beneficial or detrimental effect on the long-term course of ADHD symptoms, social–emotional functioning, motor control, timing or verbal working memory. Using a prospective longitudinal study design, we show that clinical improvement of ADHD symptoms over the course of adolescence occurs in those who are treated with stimulants during that time, as well as in those who are not.

Acknowledgement. This work was supported by National Institutes of Health Grant (R01MH62873); Netherlands Organization for Scientific Research Large Investment Grant (JB, 1750102007010); ZonMW Priority Medicines for Children Grant (PH, 113202005); grants from Radboud University Nijmegen Medical Center, University Medical Center Groningen, Accare and the VU University Amsterdam. BF is supported by a Vici grant of the Netherlands Organization for Scientific Research (016-130-669).

Conflict of interest. Dr Buitelaar has been in the past 3 years a consultant to/member of advisory board of/and/or speaker for Janssen Cilag BV, Eli Lilly and Co., Shire, Novartis, Lundbeck and Servier. He is not an employee of any of these companies, or a stock shareholder of any of these companies. He has no other financial or material support, including expert testimony, patents or royalties. Dr Franke received an educational speaking fee from Merz. Dr Hoekstra has been a paid consultant to Shire and Eli Lilly and Co. Dr Hartman, Dr Heslenfeld, Dr Rommelse, Dr Oosterlaan, Dr Schweren and Ms. van Lieshout report no potential conflicts of interest.

Ethical standards. The authors assert that all procedures contributing to this work comply with the ethical standards of the relevant national and institutional committees on human experimentation and with the Helsinki Declaration of 1975, as revised in 2008.

References

Abikoff H, Hechtman L, Klein RG, Weiss G, Fleiss K, Etcovitch J *et al.* (2004) Symptomatic improvement in children with ADHD treated with long-term methylphenidate and multimodal psychosocial treatment. *Journal of the American Academy of Child and Adolescent Psychiatry* **43**(7), 802–811.

Biederman J, Monuteaux MC, Spencer T, Wilens TE and Faraone SV (2009) Do stimulants protect against psychiatric disorders in youth with ADHD? A 10-year follow-up study. *Pediatrics* **124**(1), 71–78.

Biederman J, Petty CR, O'Connor KB, Hyder LL and Faraone SV (2012) Predictors of persistence in girls with attention-deficit/hyperactivity disorder: results from an 11-year controlled follow-up study. *Acta Psychiatrica Scandinavica* **125**(2), 147–156.

Chang Z, D'Onofrio BM, Quinn PD, Lichtenstein P and Larsson H (2016) Medication for attention-deficit/hyperactivity disorder and risk for depression: a nationwide longitudinal cohort study. *Biological Psychiatry* **80**(12), 916–922.

Charach A, Ickowicz A and Schachar R (2004) Stimulant treatment over five years: adherence, effectiveness, and adverse effects. *Journal of the American Academy of Child and Adolescent Psychiatry* **43**(5), 559–567.

Coghill DR, Seth S, Pedroso S, Usala T, Currie J and Gagliano A (2014) Effects of methylphenidate on cognitive functions in children and adolescents with attention-deficit/hyperactivity disorder: evidence from a systematic review and a meta-analysis. *Biological Psychiatry* **76**(8), 603–615.

Conners CK, Erhardt D and Sparrow AP (1999) *Conner's Adult ADHD Rating Scales: CAARS*. North Tonawanda, NY: Multi-Health Systems.

Conners CK, Sitarenios G, Parker JDA and Epstein JN (1998*a*) The Revised Conners' Parent Rating Scale (CPRS-R): factor structure, reliability, and criterion validity. *Journal of Abnormal Child Psychology* **26**(4), 257–268.

https://doi.org/10.1017/S0033291718000545 Published online by Cambridge University Press

Conners CK, Sitarenios G, Parker JDA and Epstein JN (1998*b*) Revision and restandardization of the Conners' Teacher Rating Scale (CTRS-R): factor structure, reliability, and criterion validity. *Journal of Abnormal Child Psychology* **26**(4), 279–291.

Huang Y, Wang L and Chen C (2012) Long-term neurocognitive effects of methylphenidate in patients with attention deficit hyperactivity disorder, even at drug-free status. *BMC Psychiatry* **12**, 194.

Kaiser ML, Schoemaker MM, Albaret JM and Geuze RH (2014) What is the evidence of impaired motor skills and motor control among children with attention deficit hyperactivity disorder (ADHD)? Systematic review of the literature. *Research in Developmental Disabilities* **36C**, 338–357.

Kaufman J, Birmaher B, Brent D, Rao U, Flynn C, Moreci P *et al.* (1997) Schedule for affective disorders and schizophrenia for school-age children-present and lifetime version (K-SADS-PL): initial reliability and validity data. *Journal of the American Academy of Child and Adolescent Psychiatry* **36**(7), 980–988.

Klein C, Wendling K, Huettner P, Ruder H and Peper M (2006) Intra-subject variability in attention-deficit/hyperactivity disorder. *Biological Psychiatry* **60**(10), 1088–1097.

Lichtenstein P, Halldner L, Zetterqvist J, Sjolander A, Serlachius E, Fazel S *et al.* (2012) Medication for attention-deficit/hyperactivity disorder and criminality. *New England Journal of Medicine* **367**(21), 2006–2014.

Molina BSG, Hinshaw SP, Swanson JM, Arnold LE, Vitiello B, Jensen PS *et al.* and MTA Cooperative Group (2009) The MTA at 8 years: prospective follow-up of children treated for combined-type ADHD in a multisite study. *Journal of the American Academy of Child and Adolescent Psychiatry* **48**(5), 484–500.

MTA Cooperative Group (1999) A 14-month randomized clinical trial of treatment strategies for attention-deficit/hyperactivity disorder. *Archives of General Psychiatry* **56**(12), 1073–1086.

Noreika V, Falter CM and Rubia K (2013) Timing deficits in attention-deficit/hyperactivity disorder (ADHD): evidence from neurocognitive and neuroimaging studies. *Neuropsychologia* **51**(2), 235–266.

Rubia K, Noorloos J, Smith A, Gunning B and Sergeant J (2003) Motor timing deficits in community and clinical boys with hyperactive behavior: the effect of methylphenidate on motor timing. *Journal of Abnormal Child Psychology* **31**(3), 301–313.

Schweren LJS, Hartman CA, Heslenfeld DJ, van der Meer D, Franke B, Oosterlaan J *et al.* (2015) Thinner medial temporal cortex in adolescents with attention-deficit/hyperactivity disorder and the effects of stimulants. *Journal of the American Academy of Child and Adolescent Psychiatry* **54**(8), 660–667.

Shaw P, De Rossi P, Watson B, Wharton A, Greenstein D, Raznahan A *et al.* (2014) Mapping the development of the basal ganglia in children with attention-deficit/hyperactivity disorder. *Journal of the American Academy of Child and Adolescent Psychiatry* **53**(7), 780–789.e11.

Shaw P, Sharp WS, Morrison M, Eckstrand K, Greenstein DK, Clasen LS *et al.* (2009) Psychostimulant treatment and the developing cortex in attention-deficit/hyperactivity disorder. *American Journal of Psychiatry* **166**(1), 58–63.

Swanson JM, Hinshaw SP, Arnold LE, Gibbons RD, Marcus S, Hur K *et al.* (2007) Secondary evaluations of MTA 36-month outcomes: propensity score and growth mixture model analyses. *Journal of the American Academy of Child and Adolescent Psychiatry* **46**(8), 1003–1014.

Swanson JM, Kraemer HC, Hinshaw SP, Arnold LE, Conners CK, Abikoff HB *et al.* (2001) Clinical relevance of the primary findings of the MTA: success rates based on severity of ADHD and ODD symptoms at the end of treatment. *Journal of the American Academy of Child and Adolescent Psychiatry* **40**(2), 168–179.

Tsai C, Huang Y, Wu C, Hwang F, Young K, Tsai M *et al.* (2013) Long-term effects of stimulants on neurocognitive performance of Taiwanese children with attention-deficit/hyperactivity disorder. *BMC Psychiatry* **13**, 330.

van Lieshout M, Luman M, Twisk JW, van Ewijk H, Groenman AP, Thissen AJ *et al.* (2016) A 6-year follow-up of a large European cohort of children with attention-deficit/hyperactivity disorder-combined subtype: outcomes in late adolescence and young adulthood. *European Child and Adolescent Psychiatry* **25**(9), 1007–1017.

van Widenfelt BM, Goedhart AW, Treffers PDA and Goodman R (2003) Dutch version of the strengths and difficulties questionnaire (SDQ). *European Child Adolescent Psychiatry* **12**(6), 281–289.

von Rhein D, Mennes M, van Ewijk H, Groenman AP, Zwiers MP, Oosterlaan J *et al.* (2015) The NeuroIMAGE study: a prospective phenotypic, cognitive, genetic and MRI study in children with attention-deficit/hyperactivity disorder. Design and descriptives. *European Child and Adolescent Psychiatry* **24**(3), 265–281.

Willcutt EG, Doyle AE, Nigg JT, Faraone SV and Pennington BF (2005) Validity of the executive function theory of attention-deficit/hyperactivity disorder: a meta-analytic review. *Biological Psychiatry* **57**(11), 1336–1346.

Yang L, Cao Q, Shuai L, Li H, Chan RC and Wang Y (2012) Comparative study of OROS-MPH and atomoxetine on executive function improvement in ADHD: a randomized controlled trial. *International Journal of Neuropsychopharmacology* **15**(1), 15–26.