StopWatch: Pilot study for an Apple Watch application for youth with ADHD

John E Leikauf[1] ⓘ, Carlos Correa[2], Andrew N Bueno[1],
Vicente Peris Sempere[1] and Leanne M Williams[1]

Abstract

Introduction: To address the need for non-pharmacologic, scalable approaches for managing attention-deficit and hyperactivity disorder (ADHD) in young people, we report the results of a study of an application developed for a wearable device (Apple Watch) that was designed to track movement and provide visual and haptic feedback for ADHD.

Methods: Six-week, open label pilot study with structured rating scales ADHD and semi-structured qualitative interview. Apple Watch software application given to users that uses actigraphy and graphic interface as well as haptic feedback to provide feedback to users about level of movement during periods of intentional focus. Linear mixed models to estimate trajectories.

Results: Thirty-two participants entered the study. This application was associated with improvement in ADHD symptoms over the 6 weeks of the study. We observed an ADHD-Rating Scale change of $\beta = -1.2$ units/week (95% CI $= -0.56$ to -1.88, $F = 13.4$, $P = .0004$).

Conclusions: These positive clinical outcomes highlight the promise of such wearable applications for ADHD and the need to pursue their further development.

Introduction

Attention-Deficit/Hyperactivity Disorder (ADHD) is a highly prevalent and impairing chronic condition that begins early in life, now estimated to affect 5–10% of youth.[1,2] Current pharmacologic treatments are generally effective for core symptoms of the disorder, but do not modify the longitudinal course of symptoms.[3] Furthermore, they often come with side effects that decrease tolerability and lead to significant discontinuation when taken over long periods of time.[4] Most families stop using medications within 1–3 years, citing tolerability and lack of efficacy, despite good evidence for efficacy and high response rates in short term studies.[5–7] However, ADHD symptoms usually persist to an impairing degree into adulthood even when syndromatic remission is achieved.[3,8] Non-pharmacologic

therapies have been found to be effective for many associated symptoms including anxiety, family functioning, depression, and oppositionality, but have been of limited effectiveness for core symptoms of the disorder. When available, inclusion of non-pharmacologic strategies in addition to medication is generally found to be preferable by many families.[9] However, non-pharmacologic therapies are difficult for many families to access due to costs and

---

[1]Stanford University, Stanford, CA, USA
[2]Princeton University, Princeton, NJ, USA

Corresponding author:
John E Leikauf, Stanford University, 401 Quarry Rd, Stanford, CA 94305, USA.
Email: jleikauf@stanford.edu

availability. One commonly encountered difficulty is the pervasive and chronic nature of the behaviors that result from core symptoms. These behaviors of inattention and hyperactivity[10] operate on a very short time scale and are thus very difficult for parents or caregivers to consistently monitor and modify with operant conditioning.

To our knowledge, no prior therapies have attempted to target symptoms of inattention or hyperactivity using automated biofeedback *in situ*. Context matters a great deal for learning new behaviors. Current neurofeedback as well as cognitive training approaches have been effective for improving scores on psychological testing paradigms in the laboratory but effects have not generalized to real-word settings.[11,12] There is also a substantial literature among education experts regarding the application of information and communication technology to cognitive training in youth with ADHD.[13] These technologies have predominately used computerized measures of the cognitive domains of inhibition, sustained attention, and working memory that are often observed to be impaired in youth with ADHD.[14] As reviewed in a recent meta-analysis, however, with blinded assessments cognitive training has been demonstrated to have only limited effect on ADHD symptoms and a dissociation is observed between improvements in the specific cognitive constructs targeted by the training (*e.g.,* working memory) and blinded ADHD symptom assessments.[12] The authors suggested, among other possibilities, that this may be due to a lack of generalization of the training effects to real-world settings or perhaps because the construct with the largest effects, working memory, is not as useful a clinical target in youth with ADHD as has been theorized.

Commercial wearable mobile devices that can track movement using accelerometer data and also provide different forms of feedback to users are increasingly available. Lower resolution actigraphy is well-established for measurement of ADHD symptoms with moderate to large effect sizes differentiating youth with ADHD from typically developing controls and effects of medication from placebo.[15,16] To our knowledge, there has been no prior attempt to use these technologies for treatment purposes as has been attempted with electroencephalography for neurofeedback paradigms. We therefore developed an application that can directly track and provide feedback about a core symptom of the disorder (hyperactivity) in a real-world setting such as the classroom to leverage this technology in order to address this significant gap in effective non-pharmacologic therapies for ADHD.

Methods

Study design

This study presents the results of a pilot of an Apple Watch application (aka StopWatch). The purposes of this study were to determine whether it was feasible to use this app to collect movement data (actigraphy via accelerometer), to collect concurrent clinical data using online surveys, and to conduct semi-structured exit interviews from which to derive user feedback for future versions of the app. This pilot study was open-label. The study period lasted six weeks. The application was primarily designed to collect movement data but users were able to interact with the device to select "focus sessions" during which they received gentle haptic feedback if movement exceeded 1.2 times the force of gravity. Participants were directed to use the sessions at will during this pilot in order to collect feasibility data. The protocol was approved by the Stanford University School of Medicine Institutional Review Board, and all participant-parent dyads provided verbal and written assent and informed consent. Inclusion criteria were as follows: Existing ADHD diagnosis, ability to speak and read and understand English, ability to provide assent and for caregiver with legal custody to provide informed consent, currently has or parents will purchase prior to start of study period an iPhone 5 or more recent iPhone model (the watch is inoperable without a paired iPhone). Exclusion criteria were: history of epilepsy, psychosis; any history of suicidal behaviors or current suicidal thoughts; active substance abuse or regular recreational substance or nicotine use; or significant visual or hearing impairment that would interfere with use of the device. The main outcome measure was the ADHD rating scale (ADHD-RS),[17] a well-validated and widely accepted clinical assessment completed by parents and frequently used in clinical trials. High scores represent greater (worse) symptom burden. Each of the 9 symptoms in each subscale are from the diagnostic criteria of the disorder and are rated from 0 to 3. The subscale scores range 0 to 27 and can be combined for a total score range 0 to 54. At the end of the study, study personnel conducted semi-structured exit interviews with the participant dyads using the following questions: What worked and why? What didn't and why? What features would you like to see added? What would make it more useful? What is your child's age? How often did your child actually use the App? Did you use the tracking feature on the phone?

Statistical analysis

All analyses were performed in R.[18] The tidyr, ggplot2, psych, lmer, lme4 and lmerTest packages were used for

modeling, visualization, and data management.[19–23] Because digital and online interventions often have significant drop-out and patchy data due to different engagement profiles and usage patterns, linear mixed models were constructed using the lme4 package to account for the anticipated missingness of available data.[24] We regressed a fixed effect of time on ADHD-RS specifying participant as a random effect. P-values were calculated using Satterthwaite's method. Two separate models were constructed for each of the ADHD-RS subscales.

Results

Participant characteristics

We recruited 32 participants. The caregivers of 56 young persons contacted our study coordinator, and 32 were deemed eligible by study criteria. All 32 completed the baseline clinical assessment. Participants' ages ranged 8 through 17 with a median age of 11. Fifteen identified as females and 17 as males. The mean baseline ADHD-Rating Scale (ADHD-RS) Inattentive score was 19.5 ($s = 4.8$) and the mean baseline ADHD-RS Hyperactive/Impulsive score was 13.0 ($s = 7.8$).

Engagement

All but three participants completed at least two separate assessment timepoints, 22 completed three or more, 20 completed 4 or more, and 11 completed every week. 14 participants completed the end of week six assessment. All but six participants' actigraphy data was collected and successfully uploaded to secure servers.

Symptom change

Significant improvements were seen in parent-rated ADHD-RS total score with $\beta = -1.2$ units/week and 95% CI $= -0.56$ to -1.88 ($F = 13.4$, $P = .0004$). Each subscale was also observed to improve (Figure 1). The model for ADHD-RS Inattentive scores yielded an estimate of change as follows: $\beta = -0.8$ units/week and 95% CI $= -0.41$ to -1.2 ($F = 16.9$, $P = .00007$). For ADHD-RS Hyperactive/Impulsive, the model yielded: $\beta = -0.4$ units/week and 95% CI $= -0.08$ to -0.8 ($F = 5.9$, $P = .02$). For ADHD-RS, negative scores represent improvement in clinical symptoms. Medication status, gender, and race/ethnicity were added as covariates but were nonsignificant. Age was a significant covariate added to the model as a random effect and improved the model fit ($Chi^2 = 17.5$, $P = .00003$), with older participants showing more improvement. The main treatment effect remained significant and essentially unchanged in magnitude with inclusion of age in the model ($\beta = -1.2$ units/week, $P = .0004$).

Qualitative assessment and feedback

Feedback was obtained from all but two participants at the end of the study period. Participants and their parents were asked what features they would have liked to have seen added that might improve the experience. The two most common user suggestions, each endorsed by over half of the participant pool, were to improve the visual tracking feature and to allow user adjustment of the movement threshold for haptic feedback. No other suggestions were mentioned by multiple participants. No adverse events were reported.

Discussion

We successfully piloted a first-of-its-kind application using a commercially available wearable device (Apple Watch Series 0) to track and provide feedback about ADHD symptoms and movement to users. A wireframe schematic is presented in Figure 2. The results support the feasibility of this application for use in the study of ADHD. The observed clinical improvement suggests a future role for such a novel application within ADHD treatment plans and strongly supports further research into its use.

This study was conducted primarily as a feasibility pilot in order to determine whether the application was basically functional and to elicit user feedback. The improvement in clinical scores was unexpected given the fairly rudimentary graphical interface and limited ability to customize the haptic feedback threshold. However, the theoretical justification for incorporation of feedback from target symptoms/biomarkers into non-pharmacologic therapies is strong. Many physiologic functions that are difficult and effortful to modify with conscious attention are responsive to feedback.[25] Actigraphy, as noted above, directly tracks some of the behavioral symptoms of interest. Children with ADHD are also more sensitive than others to the immediacy of rewards.[26]

Importantly, the application itself was tolerated well and the users provided clear feedback that will guide future steps. Consistently, users requested an improved graphical interface to track symptoms and that the threshold for haptic feedback could be adjusted. In fact, many participants felt the threshold was too sensitive, while some felt it could have been more sensitive. On review of the feedback we grouped these together under requests to change the haptic threshold. Our team is at work incorporating these suggestions into an improved version of the application. Other future directions include solutions for automated assent of
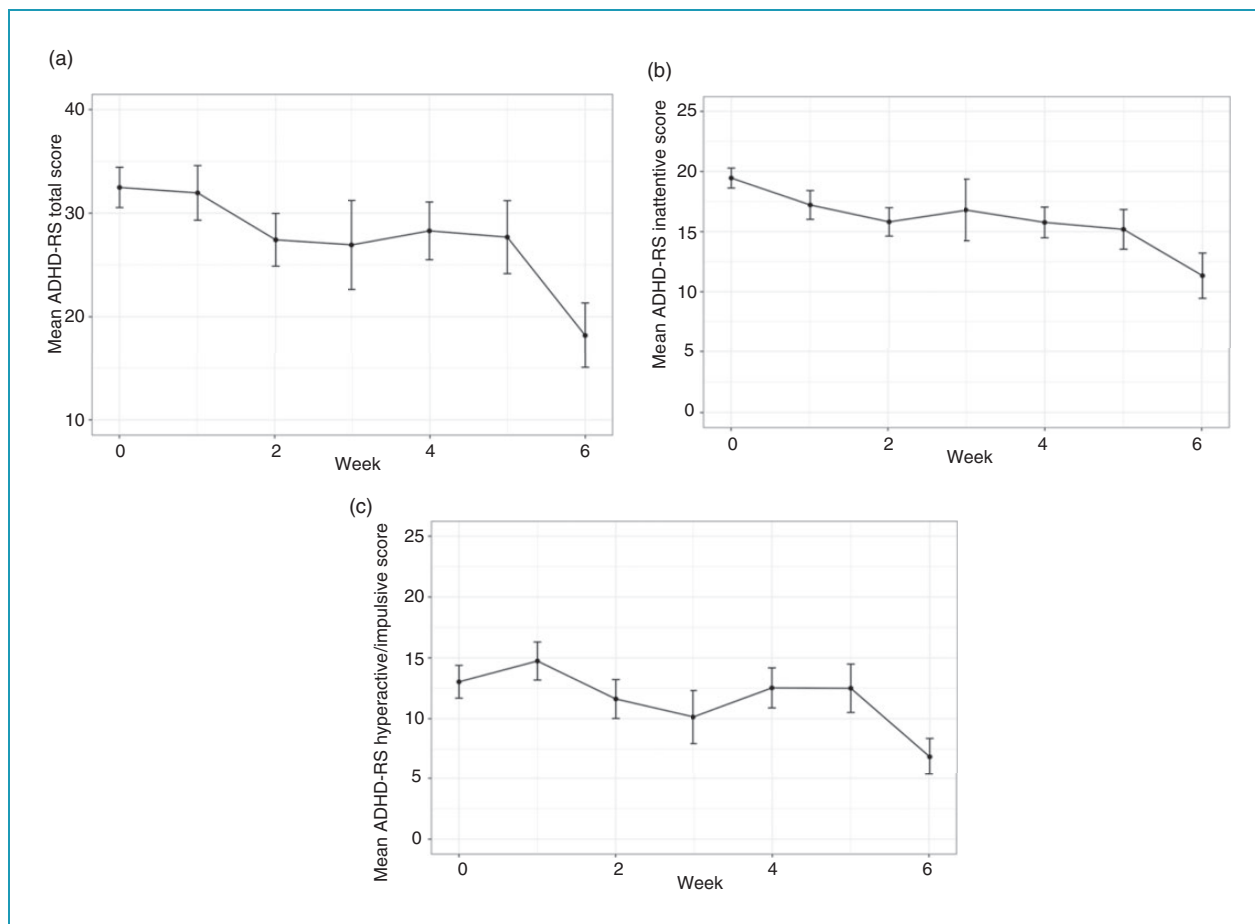
Figure 1. Plots of change in overall attention-deficit/hyperactivity disorder (ADHD) symptoms against time. ADHDRS represents the ADHD (a) Rating Scale, a gold-standard clinical outcome measure with two subcomponents representing the two domains of ADHD symptoms: (b) inattentive and (c) hyperactive/impulsive. These mirror the symptoms used for the diagnostic criteria of the disorder. Higher scores represent worse symptoms so decrease in scores represents improvement.

minors and inclusion in developmentally appropriate behavioral plans (e.g. explicit rewards programs for younger children, self-monitoring with relatively less parental involvement for teens). Targeting the intervention more specifically to different age groups was anticipated to be a future development goal and our finding of a significant effect of age is consistent with our expectation that teenagers may be able to use the app more independently while younger children would benefit more from parental scaffolding.

Limitations include those expected based on the nature of this study. Chiefly, the study was conducted open-label. Thus, the estimate of improvement is likely to be biased by user/parent expectation. However, despite the strong qualitative feedback that improvements are needed, parents still reported improvement during the course of the study. While symptom ratings are subjective, ADHD symptoms do not typically spontaneously improve to a significant degree over a period of 6 weeks. Further, medication status did not

appear to exert an effect on the results. There is no theoretical reason why the application would be less likely to be useful when combined with medication. In fact, one potential future use of the application would be to assist in tracking effects of medication changes.

Also of note, our team ran into several practical challenges in designing the application. With a limited research study budget and limited software and hardware engineering expertise in the academic research setting, it was not feasible to develop applications for multiple platforms. There is no reason in principle why this strategy could not also be applied to other devices by different manufacturers. There were also challenges to use of the Apple Watch in youth as the Watch needs to be paired with an iPhone. Most dyads elected to pair with the parent's phone. Finally, due to Apple's built-in and not publicly described strategies to limit both battery drain and apps' surreptitious use of personal data for commercial purposes, our
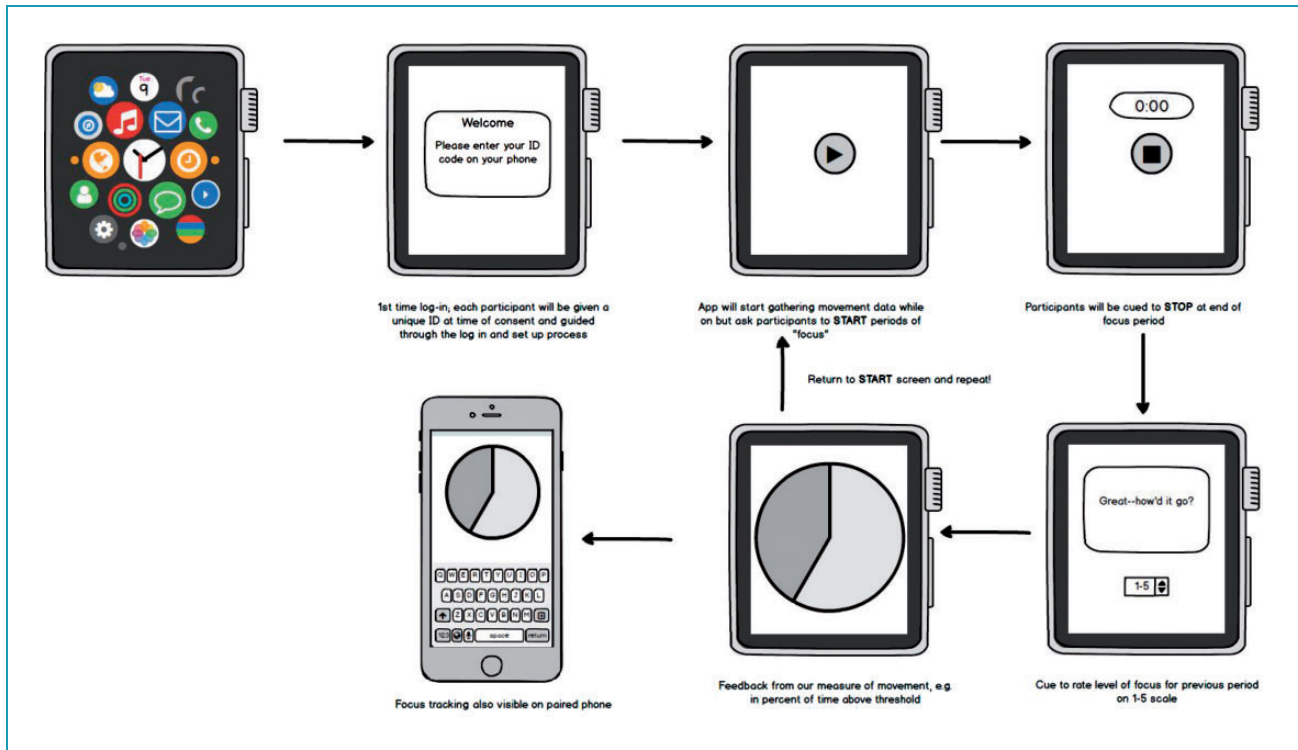
Figure 2. Wireframe of Apple Watch application. The user flows through selecting a focus session, with feedback about percentage of time spent below movement threshold (0.2 G in pilot) and opportunity for self rating. Haptic feedback is also provided when the user's movement exceeds this threshold.

programming team had to find workarounds to upload the accelerometer data reliably.

Nevertheless, the study also had important strengths. This innovative application is, to our knowledge, the first attempt to implement biofeedback for youth with ADHD *in situ*. The approach takes advantage of increasingly widely available devices that many people already own and is potentially highly scalable. Further, the improvements observed were robust to concurrent medication use and were observed to generalize across gender and ethnicity in our sample. These positive clinical outcomes highlight the promise of such wearable applications for ADHD and the need to pursue their further development.

Conclusions

We report the initial results of a pilot study for a novel application for youth with ADHD symptoms that tracks movement and provides the user haptic and visual feedback regarding their movement during self-selected periods of intentional focus. We observed improvement in ADHD symptoms for participants. There were no issues with tolerability. These findings support further development and study of this first-of-its-kind actigraphy-based biofeedback for persons with ADHD.

ORCID iD: John E Leikauf https://orcid.org/0000-0001-7364-1106

References

1. Willcutt EG, Nigg JT, Pennington BF, et al. Validity of DSM-IV attention deficit/hyperactivity disorder symptom dimensions and subtypes. *J Abnorm Psychol* 2012; 121: 991–1010.
2. Visser SN, Danielson ML, Bitsko RH, et al. Trends in the parent-report of health care provider-diagnosed and medicated attention-deficit/hyperactivity disorder: United States, 2003–2011. *J Am Acad Child Adolesc Psychiatry* 2014; 53: 34–46. e32.
3. Faraone SV, Asherson P, Banaschewski T, et al. Attention-deficit/hyperactivity disorder. *Nat Rev Dis Primers* 2015; 1: 15020.
4. Swanson JM, Arnold LE, Molina BS, et al.; the MTA Cooperative Group. Young adult outcomes in the follow-up of the multimodal treatment study of attention-deficit/hyperactivity disorder: symptom persistence, source discrepancy, and height suppression. *J Child Psychol Psychiatr* 2017; 58: 663–678.
5. Visser SN, Bitsko RH, Danielson ML, et al. Treatment of attention deficit/hyperactivity disorder among children with special health care needs. *J Pediatr* 2015; 166: 1423–1430. e1421–1422.
6. Hinshaw SP, Arnold LE and MTA Cooperative Group. Attention-deficit hyperactivity disorder, multimodal treatment, and longitudinal outcome: evidence, paradox, and challenge. *Wiley Interdiscip Rev Cogn Sci* 2015; 6: 39–52.
7. Barbaresi WJ, Katusic SK, Colligan RC, et al. Long-term stimulant medication treatment of attention-deficit/hyperactivity disorder: results from a population-based study. *J Dev Behav Pediatr* 2014; 35: 448–457.
8. Uchida M, Spencer TJ, Faraone SV, et al. Adult outcome of ADHD: an overview of results from the MGH longitudinal family studies of pediatrically and psychiatrically referred youth with and without ADHD of both sexes. *J Atten Disord* 2018; 22: 523–534.
9. Jensen PS, Hinshaw SP, Swanson JM, et al. Findings from the NIMH multimodal treatment study of ADHD (MTA): implications and applications for primary care providers. *J Dev Behav Pediatr* 2001; 22: 60–73.
10. American Psychiatric Association and American Psychiatric Association. DSM-5 Task Force. *Diagnostic and statistical manual of mental disorders: DSM-5*. 5th ed. Washington, DC: American Psychiatric Association, 2013, p.xliv, 947 p.
11. Loo SK and Makeig S. Clinical utility of EEG in attention-deficit/hyperactivity disorder: a research update. *Neurotherapeutics* 2012; 9: 569–587.
12. Cortese S, Ferrin M, Brandeis D, et al.; European ADHD Guidelines Group (EAGG). Cognitive training for attention-deficit/hyperactivity disorder: meta-analysis of clinical and neuropsychological outcomes from randomized controlled trials. *J Am Acad Child Adolesc Psychiatry* 2015; 54: 164–174.
13. Drigas A and Tourimpampa A. Processes and ICT tools for ADHD assessment, intervention and attention training. *Int J Emerg Technol Learn* 2014; 9: 20.
14. Nigg JT. Neuropsychologic theory and findings in attention-deficit/hyperactivity disorder: the state of the field and salient challenges for the coming decade. *Biol Psychiatry* 2005; 57: 1424–1435.
15. De Crescenzo F, Licchelli S, Ciabattini M, et al. The use of actigraphy in the monitoring of sleep and activity in ADHD: a meta-analysis. *Sleep Med Rev* 2016; 26: 9–20.
16. De Crescenzo F, Armando M, Mazzone L, et al. The use of actigraphy in the monitoring of methylphenidate versus placebo in ADHD: a meta-analysis. *Atten Defic Hyperact Disord* 2014; 6: 49–58.
17. DuPaul GJTJ and Anastopoulos AD. *ADHD rating scale-IV: checklists, norms and clinical interpretation*. New York, NY: Guilford Press, 1998.
18. Team RC. *R: a language and environment for statistical computing*. Vienna: R Foundation for Statistical Computing, 2019.
19. Wickham H. tidyverse: Easily Install and Load the 'Tidyverse', https://CRAN.R-project.org/package = tidy verse, 2017.
20. Wickham H. *ggplot2: elegant graphics for data analysis*. New York, NY: Springer, 2009.
21. Revelle W. *Psych: Procedures for personality and psychological research*. Evanston, IL, 2018.
22. Venables WN and Ripley BD. *Modern appliied statistics with S*. 4th ed. New York, NY: Springer, 2002.
23. Kuznetsova A, Brockhoff PB and Christensen RHB. lmerTest package: tests in linear mixed effects models. *J Stat Soft* 2017; 82: 1–26.
24. Hamer RM and Simpson PM. Last observation carried forward versus mixed models in the analysis of psychiatric clinical trials. *Am J Psychiatry* 2009; 166: 639–641.
25. Schwartz GE. Biofeedback and the behavioral treatment of disorders of disregulation. *Yale J Biol Med* 1979; 52: 581–596.
26. Stevens MC, Pearlson GD, Calhoun VD, et al. Functional neuroimaging evidence for distinct neurobiological pathways in attention-deficit/hyperactivity disorder. *Biol Psychiatry Cogn Neurosci Neuroimaging* 2018; 3: 675–685.