

## REVISIÓN ESCÉPTICA DEL ESTADO DE LA NEUROIMAGEN EN EL TRASTORNO POR DÉFICIT DE ATENCIÓN E HIPERACTIVIDAD

F. XAVIER CASTELLANOS

*Nathan Kline Institute for Psychiatric Research, Orangeburg, NY USA*

**Resumen** El trastorno por déficit de atención e hiperactividad (TDAH) ha sido estudiado por medio de resonancia magnética durante más de 30 años, superando 2200 artículos en *PubMed*. Sin embargo, todavía no se entienden bien las bases cerebrales del TDAH. Esto refleja la dificultad de replicar resultados, que afecta a casi todos los esfuerzos científicos. Los factores que contribuyen a resultados falsos positivos incluyen tamaños de muestra pequeños y la superabundancia de métodos analíticos. En el campo de la genética molecular, estos retos conllevaron a la adopción del requisito de compartir los datos genéticos inmediatamente para que la comunidad pueda trabajar de forma conjunta, y que se apliquen métodos rigurosos tomando en cuenta la verdadera cantidad de pruebas estadísticas. Esto ha producido resultados más creíbles, aunque con tamaños de efecto muy reducidos respecto a los anteriores. En este breve resumen se usan dos consorcios, uno internacional llamado ENIGMA (*Enhancing Neuro-Imaging Genetics through Meta-Analysis*), y el otro norteamericano llamado ABCD (*Adolescent and Cognitive Behavior Development Study*), para ilustrar este movimiento a la ciencia abierta. Aquí se revisa la primera cosecha de hallazgos, aunque todavía limitados a análisis transversales. Ya que ABCD fue diseñado como esfuerzo longitudinal, la combinación con la maduración continua del campo promete transformar nuestro entendimiento de la patofisiología de TDAH (probablemente también alterando la definición diagnóstica al largo plazo) para acercarnos al día en el cual la neuroimagen sea útil en la clínica.

**Palabras clave:** resonancia magnética, ciencia abierta, crisis de replicabilidad, TDAH

**Abstract** *Skeptical review of the state of neuroimaging in attention deficit hyperactivity disorder.* Attention-deficit/hyperactivity disorder (ADHD) has been the focus of magnetic resonance imaging studies for more than 30 years, with more than 2200 articles listed in PubMed. Nevertheless, the brain substrates of ADHD remain poorly understood. This reflects the crisis of replicability across nearly all scientific endeavors, deriving from factors such as small sample sizes combined with a proliferation in analytical approaches, yielding high rates of false positive results. The field of molecular genetics confronted this by adopting open and immediate sharing of raw data and insistence on rigorous corrections for multiple comparisons. These strategies are yielding more robust genetic findings, albeit with much smaller effect sizes than before. This brief review focuses on two recent consortium efforts, i.e., the international Enhancing Neuro-Imaging Genetics through Meta-Analysis (ENIGMA), and the U.S. Adolescent Behavior & Cognitive Development Study (ABCD). Both embrace the culture of open science, and are beginning to yield credible findings, despite being limited initially to cross-sectional analyses. As the field continues to mature, these and other ongoing longitudinal large-scale studies are poised to transform our understanding of the pathophysiology of ADHD to bring closer the day when neuroimaging can contribute to clinical utility.

**Key words:** magnetic resonance imaging, open science, crisis of replicability, ADHD

El trastorno por déficit de atención e hiperactividad (TDAH) ha sido estudiado por medio de resonancia magnética (MRI) durante más de 30 años, superando 2200 artículos (búsqueda sencilla de ADHD + MRI en *PubMed*) a finales del 2021. Sin embargo, todavía no se entienden bien las bases cerebrales del TDAH. En este breve resumen, se destacan los hallazgos que parecen ser más replicables, los cuales sugieren las direcciones que se debieran seguir antes de que se pueda incorporar

estudios del cerebro en la clínica. Para economizar el espacio, el lector será referido a una revisión reciente del campo de neuroimagen<sup>1</sup> y a las “208 conclusiones basadas en datos” recientemente compiladas en una declaración conjunta de la Federación Mundial del TDAH<sup>2</sup>.

### Respondiendo a la crisis de replicabilidad

Para comenzar, es notable que dentro de las 208 conclusiones basadas en datos sobre TDAH, solo cinco tratan de estudios de MRI<sup>2</sup>. La primera de éstas proviene de estudios llevados a cabo por el proyecto colaborativo llamado ENIGMA (del inglés *Enhancing Neuro-Imaging*

**Dirección postal:** F. Xavier Castellanos, MD, NYU Grossman School of Medicine. 1 Park Avenue, 7th Floor, New York, NY USA 10016  
e-mail: Francisco.castellanos@nyulangone.org

*Genetics through Meta Analyses*). Como indica el nombre, este esfuerzo fue lanzado con el propósito de conectar estudios genéticos con los de MRI. En la genética, el campo empezó con optimismo y la publicación de muchos hallazgos que no fueron replicados. En respuesta, las agencias principales que financian la mayoría de los estudios de genética insistieron en la práctica de *open science*, requiriendo que todos los laboratorios apoyados por ellas compartieran datos brutos desde el inicio, sin esperar hasta que se publicaran<sup>3</sup>. Eso lanzó la época de muestras grandísimas (*Big Data*) y métodos estadísticos controlando rigurosamente el riesgo de falsos positivos (error tipo I) para poder detectar señales replicables en estudios del genoma entero (*genome-wide association studies* = GWAS). De hecho, en TDAH, después de muchos estudios de genes candidatos, no se lograron hallazgos estadísticamente creíbles en estudios GWAS hasta que se juntó una muestra de más de 20 000 pacientes diagnosticados con TDAH y más de 35 000 controles sanos<sup>2</sup>. Esa primera docena de hallazgos GWAS representa el comienzo de una nueva etapa en la genética del TDAH, pero solo indica pistas de que algunas variaciones genéticas en el vecindario de los hallazgos aportan un aumento de riesgo de manifestar TDAH. No identifican los genes involucrados, y los incrementos de riesgo son pequeños<sup>2</sup>. Estas lecciones son relevantes para la neurociencia clínica y especialmente para el campo de MRI.

Agregar datos cerebrales obtenidos por MRI es más complicado que hacerlo para estudios genéticos. El genoma es compuesto de solo cuatro "letras" o bases, y todos los métodos analíticos debieran producir el mismo código genético, por lo menos en teoría. Si dos métodos divergen, la muestra puede ser casi perfectamente duplicada y analizada en múltiples centros, hasta que se llegue a resultados de consenso. En cuanto a la medición del cerebro, hay múltiples niveles de complejidad. Primero, la estructura fundamental del genoma es casi idéntica en todos. Los cerebros típicos también comparten muchas propiedades, pero cada cerebro es tan individual como las huellas digitales. El cerebro cambia a través de la vida, especialmente durante la niñez y la vejez, y por lo menos en cuanto a tamaño, las diferencias más notables son entre personas masculinas y femeninas. Pasando al próximo nivel de divergencia, existen solo tres compañías en el mundo que fabrican resonadores de alto campo magnético, pero cada una tiene varios modelos, por ejemplo, con diámetros más o menos grandes, campos magnéticos de 1.5 Tesla, 3.0 Tesla (actualmente el más común usado en investigaciones) o 7 Tesla, lo cual también afecta la calidad y precisión de los resultados. De hecho, todas variaciones, aunque menores, pueden influenciar los resultados, pero no es factible que todas las muestras sean estudiadas en la misma máquina MRI y que esa no se actualice en cuanto a *software*, lo cual típicamente afecta los resultados de manera ligera

pero complicada. Así que en este campo se están desarrollando métodos de ajuste estadístico por los efectos asociados con obtener datos MRI en un centro dado, que comprenden variaciones en *hardware*, *software*, y factores todavía desconocidos<sup>4</sup>. Una vez que los datos estructurales y funcionales han sido obtenidos y pre-procesados, existen múltiples maneras de medirlos, cada una con sus ventajas y dificultades. La buena noticia es que, en cuanto a datos estructurales, estos retos están siendo superados, resultando en valores de fiabilidad (*test-retest reliability*) que típicamente exceden 0.95<sup>5</sup>.

La medición de la estructura cerebral se ha beneficiado dramáticamente por la adopción de un sistema de *software* abierto, llamado *FreeSurfer* <https://surfer.nmr.mgh.harvard.edu/>. Se puede aplicar a cualesquiera datos cerebrales, pero su uso más frecuente es para procesar datos estructurales y rendir índices volumétricos del cerebro, usando la parcelación cerebral del atlas inicialmente producido por Desikan, Killiany y colegas<sup>6</sup>, que divide la corteza cerebral en 34 regiones por hemisferio. Para cada una de estas regiones, *FreeSurfer* da el promedio del grosor de la corteza y el tamaño de su superficie, con el volumen, siendo su multiplicación. También provee índices volumétricos de los ganglios basales (caudado, putamen, núcleo *accumbens*, pálido) y de la amígdala e hipocampo. *FreeSurfer* rinde otros índices adicionales, pero estos son los más frecuentemente usados. La idea fundamental y genial de ENIGMA fue que, aunque puede ser complicado compartir datos individuales coleccionados en distintos centros, cuando el consentimiento para agregar no fue obtenida desde el inicio, es factible compartir los índices derivados de *FreeSurfer*, ya que los índices de superficie, volumen o grosor de la corteza no representan datos personales que puedan identificar personas particulares. De esta manera, se pueden agregar muestras de cualquier tamaño para pruebas estadísticas con datos de cada individuo, llamado mega-análisis, traspasando la limitación de solo poder sintetizar medios y desviaciones estándar de los distintos estudios (meta-análisis).

### Estudios de ENIGMA sobre TDAH

El primer mega-análisis del consorcio ENIGMA en TDAH se enfocó en las estructuras subcorticales y en el volumen intracraneal<sup>7</sup>. Los datos procesados por *FreeSurfer* fueron agregados desde 23 centros. La muestra incluyó 1713 participantes con TDAH y 1520 controles sanos; la edad media era 14 años (rango 4 a 63). En el examen inicial de todos pacientes *versus* todos los controles, sin importar edades, se encontraron valores menores de los volúmenes subcorticales en el grupo TDAH, fuertemente significativos estadísticamente ( $p < 0.001$ , corregido tomando en cuenta el número de pruebas usando el método *false discovery rate*), pero con tamaños de efecto pequeños

(definido por Cohen como un valor  $d \leq 0.20$ ). En el caso del volumen intracraneal, se demostró un valor menor en el grupo TDAH con  $d = -0.10$ . Los subsecuentes análisis fueron ajustados para tomar en cuenta esa diferencia global por medio de análisis de covariables, y se encontraron tamaños de efecto  $d = -0.15$  para el núcleo *accumbens*;  $d = -0.11$  para el caudato;  $d = -0.14$  para el putamen;  $d = -0.19$  para la amígdala;  $d = -0.12$  en el hipocampo. No se encontraron diferencias estadísticas en el pálido ni en el tálamo. Los efectos fueron más fuertes en niños (o sea, menores a 14 años) que en adultos, y de hecho, ningunos de los resultados en los adultos fueron significativos. En los adolescentes (edades entre 15 y 21 años) solo el hipocampo demostró una diferencia ( $d = -0.24$ ) entre TDAH y controles.

El segundo esfuerzo sobre TDAH por el consorcio ENIGMA se enfocó en la corteza cerebral, con una muestra derivada de 36 centros internacionales, incluyendo 2246 participantes con TDAH y 1934 controles sanos<sup>8</sup>. En esta muestra ampliada, se clarificó que el tamaño menor del cerebro en TDAH es mayormente asociado con valores menores de superficie, más que del grosor de la corteza. La superficie global fue menor ( $d = -0.21$ ) en el grupo con TDAH, con valores particularmente notablemente distintos en áreas frontales, relacionadas con el cíngulo, y en áreas temporales (con valores  $d$  entre  $-0.19$  y  $-0.10$ ). Hubo diferencias en grosor de la corteza solamente en el polo temporal y en el giro fusiforme. Además, ningunas de estas diferencias fueron encontradas en adolescentes o adultos, destacando nuevamente que los hallazgos anatómicos varían según la edad<sup>8</sup>. Los autores opinaron que el hecho de que las diferencias anatómicas entre grupos se encuentran únicamente en niños, con una sola excepción, apoya la hipótesis de que el TDAH involucra un retraso en procesos fisiológicos del desarrollo, con tales efectos disminuyendo con la madurez. Sin embargo, la mayor limitación de este tipo de estudio transversal es que no puede distinguir entre distintas trayectorias en desarrollo cerebral, lo cual requiere estudios longitudinales<sup>9</sup>.

### Estudios de resonancia funcional

En cuanto a estudios funcionales, el campo está aún en peores condiciones. En 2005, Ioannidis nos advirtió que la mayoría de los resultados científicos publicados son falsos<sup>10</sup>. Sus críticas aplican particularmente a estudios de MRI funcionales, que suelen tener muestras aún más reducidas que los estudios estructurales, lo cual reduce aún más el poder estadístico<sup>11</sup>, combinado con la urgencia de obtener fondos, la popularidad del campo, y la gran diversidad en modos de analizar datos<sup>12</sup>. Todos estos factores resultan en hallazgos que no pueden ser reproducidos. Sí se han publicado varios meta-análisis que documentan diferencias en funcionamiento en TDAH,

usando MRI funcional (fMRI)<sup>1</sup>. Desafortunadamente, por lo menos algunas de los anteriores meta-análisis estuvieron afectados por errores en los algoritmos, que rindieron resultados positivos falsos, y la gran mayoría de los estudios incluidos siguen teniendo tamaños de muestra reducidos, posiblemente explicando porque el mayor meta-análisis, que incluyó 96 estudios, tanto funcionales como estructurales, no obtuvo ningún efecto significativo<sup>13</sup>.

### Resultados iniciales del estudio ABCD (Adolescent Brain & Cognitive Development Study)

¿Entonces, donde quedamos? La oportunidad de estudiar el funcionamiento del cerebro no se puede ignorar, pero hay que enfrentar con ojos abiertos los retos involucrados. Nos convencimos de que íbamos a encontrar efectos grandes, pero a medida que los métodos mejoran, nos estamos dando cuenta que casi todos los efectos verdaderos (replicables) son pequeños<sup>14</sup>. Por lo tanto, esfuerzos serios tendrán que utilizar muestras grandes. Un ejemplo es el estudio ABCD (*Adolescent Brain & Cognitive Development Study*) <https://abcdstudy.org/>, que reclutó más de 11 000 jóvenes cuando tenían 9 o 10 años con la intención de seguirlos estrechamente y con resonancias cada dos años, hasta que llegaran a los 20 años<sup>15</sup>. ABCD es un ejemplo de ciencia abierta prospectiva, ya que las familias han dado consentimiento para que sus datos, una vez pulidos de nombres e información fácilmente identificable, puedan ser compartidos con investigadores legítimos en cualquier parte del mundo. Incluye grandes cantidades de medidas tomadas de los padres, los jóvenes, sus profesores, con datos MRI estructurales y funcionales, más datos genéticos. Utilizando el método dimensional con la escala de "problemas de atención", de la escala *Child Behavior Checklist* (que no obstante su nombre, refleja TDAH y no solamente problemas de inatención) en lugar de presencia o ausencia de TDAH, Owens y colegas exploraron la posibilidad de encontrar huellas neurobiológicas relacionadas con tal escala<sup>16</sup>. Enfocándose en MRI funcional con tareas de memoria de trabajo, inhibición, y recompensa, en combinación con MRI estructural, pudieron encontrar una relación estadísticamente significativa usando la tarea de memoria de trabajo, pero no las otras. Aunque fue fuertemente significativa estadísticamente, el valor  $R^2$  fue solo 2%. Los datos estructurales no contribuyeron de manera importante<sup>16</sup>.

Los estudios funcionales que usan MRI tienen que enfocarse en alguna tarea, por ejemplo, memoria de trabajo, con el objetivo de identificar las regiones más fuertemente asociadas con la ejecución de la tarea. Al otro extremo, se pueden estudiar las fluctuaciones espontáneas de la señal BOLD (*blood oxygen level dependent*) sin tarea alguna más de quedarse en el *scanner* por 5 a 20 minutos, preferiblemente con los ojos abiertos<sup>1</sup>. Este método,

llamado fMRI *en estado de reposo*, ofrece la posibilidad de revelar las grandes redes neuronales intrínsecas del cerebro. La desventaja mayor de este método es su extrema sensibilidad a los movimientos de la cabeza, lo cual es problemático para todo tipo de estudio MRI, pero particularmente para estudios en estado de reposo<sup>1</sup>. El estudio ABCD ha implementado procedimientos para intentar minimizar los efectos de tales movimientos, pero también ofrece una muestra suficientemente grande para permitir excluir datos que contengan demasiado movimiento. En un trabajo inicial, Karcher y colegas examinaron índices de conectividad funcional entre grandes redes neuronales e índices de psicopatología global (el 'factor P') y un sistema alternativo de taxonomía de psicopatología (HITOP, *Hierarchical Taxonomy of Psychopathology*), adaptado para jóvenes, que rindió entre 2 y 5 factores, incluyendo un factor de neurodesarrollo, que comprende subfactores de inatención, hiperactividad, impulsividad, torpeza, y compartimientos repetitivos<sup>17</sup>. Dado al tamaño de la muestra ABCD, los investigadores pudieron dividirla en dos, aleatoriamente. Una, con n= 3790, fue siempre usada para análisis iniciales, y fue llamada la muestra de descubrimiento. Los investigadores solamente reportaron hallazgos que fueron replicados en la segunda muestra (n = 3791) dedicada a la replicación. El factor de neurodesarrollo, que se superpone en gran parte con TDAH<sup>17</sup>, demostró asociaciones en ambas muestras con niveles más bajos de conectividad funcional entre la red por defecto (DMN, del inglés *default mode network*) con sí misma, y entre la DMN y la red cíngulo-opercular<sup>18</sup>.

### ¿A dónde vamos?

Estos resultados tendrán que ser comprobados longitudinalmente en subsiguientes fases del estudio ABCD, a medida que los jóvenes maduran. Los factores HITOP tanto como el factor P son fácilmente calculados de una escala llenada por padres (*Child Behavior Checklist*) que ha sido traducida a 110 lenguajes, incluyendo versiones en español para Argentina, Chile, México, España y para hispanohablantes en EE.UU., <https://aseba.org/translations/>. Es verdad que el sistema Achenbach es propietario, lo cual representa un obstáculo para su uso universal, pero su costo es mínimo en comparación con el gasto de la resonancia. Si estudios independientes validan la versión pediátrica de la taxonomía HITOP<sup>17</sup>, podría reemplazar los sistemas actuales del DSM-5 e ICD-11 en la investigación, con el propósito de mejorar nuestro entendimiento de las bases neurobiológicas de los trastornos psiquiátricos como TDAH. Sin embargo, no debiéramos pensar que métodos como MRI reemplazarán la evaluación clínica en cuanto a establecer la existencia de psicopatología y la necesidad de iniciar tratamientos. Lo que deseamos es que las mejoras en métodos, tanto

de definir fenotipos como de medir procesos fisiológicos neuronales nos aporten entendimientos de la fisiología de condiciones como TDAH, probablemente transformando nuestros conceptos diagnósticos en el proceso, para llegar algún día al sueño de la medicina de precisión<sup>19</sup>.

**Conflictos de interés:** El autor sirve en la junta de asesoría de BOL Pharma, lo cual no está relacionado con este trabajo.

### Bibliografía

1. Pereira-Sanchez V, Castellanos FX. Neuroimaging in attention-deficit/hyperactivity disorder. *Current Opinion in Psychiatry* 2021; 34: 105-11.
2. Faraone SV, Banaschewski T, Coghill D, et al. The World Federation of ADHD International Consensus Statement: 208 Evidence-based conclusions about the disorder. *Neuroscience and Biobehavioral Reviews* 2021; 128: 789-818.
3. Maxson Jones K, Ankeny RA, Cook-Deegan R. The Bermuda Triangle: The Pragmatics, Policies, and Principles for Data Sharing in the History of the Human Genome Project. *Journal of the History of Biology* 2018; 51: 693-805.
4. Radua J, Vieta E, Shinohara R, et al. Increased power by harmonizing structural MRI site differences with the ComBat batch adjustment method in ENIGMA. *NeuroImage* 2020; 218:116956. doi:10.1016/J.NEUROIMAGE.2020.116956
5. Rezende TJR, Campos BM, Hsu J, et al. Test-retest reproducibility of a multi-atlas automated segmentation tool on multimodality brain MRI. *Brain and Behavior* 2019; 9(10). doi:10.1002/BRB3.1363
6. Desikan RS, Ségonne F, Fischl B, et al. An automated labeling system for subdividing the human cerebral cortex on MRI scans into gyral based regions of interest. *NeuroImage* 2006; 31: 968-80.
7. Hoogman M, Bralten J, Hibar DP, et al. Subcortical brain volume differences in participants with attention deficit hyperactivity disorder in children and adults: a cross-sectional mega-analysis. *The Lancet Psychiatry* 2017; 4: 310-319.
8. Hoogman M, Muetzel R, Guimaraes JP, et al. Brain imaging of the cortex in ADHD: A coordinated analysis of large-scale clinical and population-based samples. *American Journal of Psychiatry* 2019; 176: 531-42.
9. Kraemer HC, Yesavage JA, Taylor JL, Kupfer D. How can we learn about developmental processes from cross-sectional studies, or can we? *American Journal of Psychiatry* 2000; 157:163-71.
10. Ioannidis JPA. Why most published research findings are false. *PLoS Medicine* 2005; 2(8): 2-8.
11. Button KS, Ioannidis JPA, Mokrysz C, et al. Power failure: Why small sample size undermines the reliability of neuroscience. *Nature Reviews Neuroscience* 2013; 14: 365-76.
12. Botvinik-Nezer R, Holzmeister F, Camerer CF, et al. Variability in the analysis of a single neuroimaging dataset by many teams. *Nature* 2020; 582(7810): 84-8.
13. Samea F, Soluki S, Nejati V, et al. Brain alterations in children/adolescents with ADHD revisited: A neuroimaging meta-analysis of 96 structural and functional studies. *Neuroscience and Biobehavioral Reviews* 2019; 100: 1-8.
14. Owens MM, Potter A, Hyatt CS, et al. Recalibrating

- expectations about effect size: A multi-method survey of effect sizes in the ABCD study. *PLOS ONE* 2021; 16(9): e0257535. doi:10.1371/JOURNAL.PONE.0257535
15. Garavan H, Bartsch H, Conway K, et al. Recruiting the ABCD sample: Design considerations and procedures. *Developmental Cognitive Neuroscience* 2018; 32:16-22.
  16. Owens MM, Allgaier N, Hahn S, et al. Multimethod investigation of the neurobiological basis of ADHD symptomatology in children aged 9-10: baseline data from the ABCD study. *Translational Psychiatry* 2021; 11(1). doi:10.1038/S41398-020-01192-8
  17. Michelini G, Barch DM, Tian Y, Watson D, Klein DN, Kotov R. Delineating and validating higher-order dimensions of psychopathology in the Adolescent Brain Cognitive Development (ABCD) study. *Translational Psychiatry* 2019; 9(1). doi:10.1038/S41398-019-0593-4
  18. Karcher NR, Michelini G, Kotov R, Barch DM. Associations between resting-state functional connectivity and a hierarchical dimensional structure of psychopathology in middle childhood. *Biological Psychiatry: Cognitive Neuroscience and Neuroimaging* 2021; 6(5):508-17.
  19. Fernandes BS, Williams LM, Steiner J, Leboyer M, Carvalho AF, Berk M. The new field of "precision psychiatry." *BMC Medicine* 2017; 15: 1-7.