***Commentary: Can medication effects be determined using national registry data? A cautionary reflection on risk of bias in “big data” analytics; based on Chang et al., 2016.***

**Edmund JS Sonuga-Barke, PhD\***

**Department of Psychology**

**University of Southampton**

**\*correspondence: University of Southampton, Southampton, S017 1BJ, UK.** **ejb3@soton.ac.uk**

**Word count – 1515.**

Stimulant and non-stimulant medications are recommended as part of ADHD treatment, based on the best available evidence of their costs and benefits(1). At the same time, some important clinical questions remain to be definitively resolved. This is because of both limitations in the scope of prior trials and, in some cases, their methodological and design weaknesses. Added to this, a degree of scepticism exists, in some quarters, concerning the veracity of the ADHD medication evidence-base as a whole, because of the potentially distorting effect of pharmaceutical industry involvement in trials. This has contributed to an atmosphere of doubt within the public square around the issue of ADHD medication (2). There is an urgent need for independent trials to address gaps in the evidence base and answer outstanding questions.

 For instance, there remains uncertainty over the long-term effects of ADHD medications. While the best available long-term surveillance data supports their safety(3), questions about the balance between long-term costs and benefits persist. On the face of it, this seems strange given the clinical importance of this issue and the vastness of the ADHD research enterprise. However, the conduct of extended randomised controlled trials (RCT), necessary to provide definitive evidence of long-term effects, is constrained by ethical and practical factors. Specifically, it is, unethical to withhold treatments, of known short-term efficacy, over the long term to the detriment of a patients wellbeing and; impracticable to expect patients assigned to a long term no-treatment control arm, not to seek out medication available locally as part of routine care – making control arm contamination inevitable.

 In the light of these constraints, different strategies have been adopted to gain traction on long-term medication effects. The Multi-modal Treatment of ADHD (MTA) study, for instance, combined two of these strategies(4). Its first phase was an RCT that demonstrated the clinical benefits of receiving medication according to a rigorous, evidence-based, algorithm over a 14-month period compared with the same duration of normal community care (which also often involved medication(4). In its second phase it resolved to a naturalistic longitudinal observational study when the rigorous medication protocol was suspended but longitudinal clinical outcome and medication exposure data continued to be gathered through to adulthood(5). In this phase, extended medication exposure was not associated with symptom improvement(6).

 Chang et al(7)employ a variant of an observational design - registery-based pharmaco-epidemiology - to examine whether ADHD medication alters the long-term risk of depression in individuals with ADHD. Using sophisticated survival modelling of longitudinal data from their integrated national registers, the authors find, both, significantly lower rates of diagnosed depression in ADHD patients previously prescribed stimulant medication, and a temporal correlation between medication use and depression onset and remission. The implication being that, rather than being a risk factor for the development of depression, as some have suggested, medication exposure, may be protective. As discussed by the authors this finding has biological plausibility and is potentially very important clinically, given how frequently depression develops in individuals with ADHD during adolescence and adulthood. However, it needs to carefully judged against current methodological standards for evidence generation in medicine – standards developed, precisely to reduce “risk of bias” (RoB) inherent in observational studies.

 RoB arises from a number of sources in observational studies. Chang et al(7) highlight some of these in their discussion, but a number of others also need consideration. RoB increases when diagnostic and outcome data has questionable validity. In this regard, there are generic concerns about the psychometric properties of clinical diagnoses. Is such a register-recorded diagnosis a reliable and valid indicator of the presence of a disorder? Clearly, in this regard, the somewhat arbitrary nature of referral and diagnostic processes and practices presents problems –with many individuals in a population with ADHD not receiving a diagnosis and some with a diagnosis not having ADHD(8). Such concerns, relating to the Swedish registers in particular, have been addressed, at least in part, in validation studies mentioned by the authors. Concerns about national registry data validity inevitably dovetail with the issue of cultural specificity - given that the former is tied, so closely, to international variations in clinical practice. Interestingly, related to this, the authors highlight the representative nature of the register as a study strength. However, it remains to be shown that individuals diagnosed with ADHD in the registry are representative of the actual ADHD sample in the population.

 Lack of random assignment to treatment further exacerbates RoB. As recognised by the authors, this is an inherent limitation of all observational studies of medication effects. In these, medication exposure is inevitably confounded with the complex set of factors that drive clinicians’ prescribing decisions(9). Patients’ characteristics are likely to be influential in this: It is possible, for instance, that patients with more severe and complex presentations are especially likely to be prescribed medication. Variations in the attitudes and circumstances of clinicians, patients and (in the case of child patients) their parents will likely moderate such effects. A number of statistical methods are used in an attempt to address this problem. Propensity score matching, for instance, attempts to control for factors that predict treatment allocation. However, in themselves, *ex post facto* statistical analyses can only take us so far. First, only a limited subset of potential confounders are ever measured in studies and so statistical control of is only ever partial. Second, even if the specific relation between medication exposure and outcome could be isolated statistically, a causal inference cannot be draw definitively without experimental evidence.

 The failure to mask treatment assignment from individuals involved in studies also increases RoB. Masking protects a study from the myriad biases that might arise from the treatment preferences or preconceptions of patients, therapists and/or researchers. Clearly, in register-based studies such masking is not possible. Indeed, leaving aside RoB related to parents’ and patients’ awareness of treatment assignment, it is possible, even probable, in the current study, that the same clinician gave the ADHD diagnosis, provided the treatment and made the outcome assessment (i.e., diagnosis of depression).

 Finally, the possibility of selective reporting of findings increases RoB: On the one hand, failed trials may not be reported at all (the classical file draw problem). On the other, even when studies are published findings may be “cherry picked” for reporting to prioritise statistically significant or favoured results over statistically non-significant. This risk is greatest when data-driven analyses exploit large and rich data sets to selectively pin-point and report significant effects of interest – especially in a publishing culture where non-significant effects rarely see the light of day. This form of bias is inherent to the research process, driven, as it is, by structural factors and unconscious processes. To counter it researchers undertaking RCTs are now obliged to publically pre-register study protocols, setting out the rationale for the study, the key research questions, the inclusion and exclusion criteria, the primary outcome variables and the statistical analysis plan. There is a move to extend such an approach to all studies in psychology and psychiatry(10). Such a development would be particularly valuable in reducing this form of RoB intrinsic to registry-based research.

 In summary, national registry studies, in general, and the previously published work of this internationally renown research group, in particular, have provided significant insights into the causes and consequences of psychiatric disorders. However, particular caution is needed when using this resource to draw inferences about treatment effects. This is because the sort of data limitations and design shortcomings discussed above, lead to substantial RoB. This renders the contribution of reported findings to the treatment evidence-base difficult to gauge. It might be argued, that when there is a lack of good quality trial data to address a vital clinical question, as is the case here with long-term medication effects, registry-based studies at least provide us with some information on which to base practice. However, the truth is that without knowing precisely which biases are operating and in which ways, one is left to speculate on whether reported findings represent an under- or an over-estimate of the actual medication effect, or indeed are completely spurious.

 There is currently great enthusiasm for big data in psychiatry and psychology. Indeed it brings with it many opportunities. However, when considering its strengths it is important we avoid falling into the trap of what might be called the “big data” fallacy – believing that, as Charity does to sin, “big data” in and of itself can “cover” a multitude of inherent data and design limitations; or, put another way, that RoB created by intrinsic design limitations in observational studies in some way is decreased as sample size increases. Unfortunately, because sample size/statistical power and the methodological rigour are not tradable study qualities – it can’t. It is perhaps sobering that the two great strengths of Chang et al (2016), its massive sample (data from 39,000 ADHD patients aged between 11 as 49 years was analysed) and the brilliance of the research team, who optimized every aspect of data analysis and presentation, better illustrates the inherent limitations, rather than the scientific potential, of registry-based pharmaco-epidemiology.

**References**

1. Alamañac B, Nutt DJ, Adamou M, Asherson P, Bazire S, Coghill D, Heal D, Müller U, Nash J, Santosh P, Sayal K, Sonuga-Barke E, Young SJ (2014). Evidence-based guidelines for the pharmacological management of attention deficit hyperactivity disorder: Update on recommendations from the British Association for Psychopharmacology. *Journal of Psychopharmacology, 28*, 179-203.

2. Banaschewski T, Buitelaar J, Chui CSL, Coghill D, Cortese S, Simonoff E, Wong ICK (in press) Methylphenidate for ADHD in children and adolescents: throwing the baby out with the bathwater. *Evidence Based Mental Health,*

3. Graham J, Banaschewski T, Buitelaar J, Coghill D, Danckaerts M, Dittmann RW, Doepfner M, Hamilton R, Hollis C, Holtmann M, Hulpke-Wette M, Lecendreux M, Rosenthal E, Rothenberger A, Santosh P, Sergeant J, Simonoff E, Sonuga-Barke E, Wong ICK, Zuddas A, Steinhausen HC, Taylor E (2011). European Guidelines on Managing Adverse Effects of Medication for ADHD. *European Child and Adolescent Psychiatry, 20*, 17-37.

4. MTA Cooperative Group (1999). A 14-month randomized clinical trial of treatment strategies for attention-deficit/hyperactivity disorder. *Archives of General Psychiatry, 56,* 1073-86.

5. Sibley MH, Swanson JM, Arnold LE, Hechtman LT, Owens EB, Stehli A, Abikoff H, Hinshaw SP, Molina B, Mitchell JT, Jensen P, Howard AL, Lakes KD, Pelham WE, and the MTA Cooperative Group (in press). Defining ADHD symptom persistence in adulthood: optimizing sensitivity and specificity, *Journal of Child Psychology & Psychiatry,*

6. Swanson J; Arnold LE; Molina B; Sibley M; Hechtman L; Hinshaw S; Abikoff H; Stehli A; Greenhill L; Kraemer H (in press). Young Adult Outcomes in the Follow-up of the Multimodal Treatment Study of Attention-Deficit/Hyperactivity Disorder: Symptom Persistence, Source Discrepancy, and Height Suppression. *Journal of Child Psychology & Psychiatry*

7. Chang Z, D’Onofrio BM, Quinn PD, Lichtenstein P, Larsson H (2016): Medication for attention-deficit/hyperactivity disorder and risk for depression: A nationwide longitudinal cohort study. Biol Psychiatry XX:XXX-XXX.

8. Wright N, Moldavsky M, Schneider J, Chakrabarti I, Coates J, Daley D, Kochhar P, Mills J,

Sorour W, Sayal K (2015). Practitioner Review: Pathways to care for ADHD – a systematic review of barriers and facilitators. *Journal of Child Psychology and Psychiatry, 56,* 598–617.

9. Kovshoff H, Williams S, Vrijens M, Danckaerts M, Thompson M, Yardley L, Hodgkins P, Songa-Barke EJS (2012). The decisions regarding ADHD management (DRAMa) study: uncertainties and complexities in assessment, diagnosisand treatment, from the clinician’s point of view. *European Child and Adolescent Psychiatry, 21*, 87–99.

10. Rutter M, Pickles A (2016). Annual Research Review: Threats to the validity of child psychiatry and psychology*. Journal of Child Psychology & Psychiatry, 57,* 398–416.

**Acknowledgements**

I would like to thank Dr Sam Wass for helpful comments on an early draft.

**Disclosures**

Over the last three years the author has held visiting chairs at Ghent University, Aarhus University and University of Sussex. He has received speaker fees, consultancy, research funding and conference support from Shire Pharma, speaker fees from Janssen Cilag, consultancy from Neurotech solutions, Aarhus University, Copenhagen University and Berhanderling, Skolerne, Copenhagen, KU Leuven. Book royalties from OUP and Jessica Kingsley. He is editor-in-chief of the Journal of Child Psychology and Psychiatry, from whom he receives an honorarium.