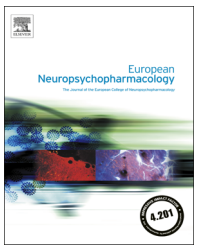
European Neuropsychopharmacology (2013) 23, 1307–1309

www.elsevier.com/locate/euroneuro



SHORT COMMUNICATION

Why negative meta-analyses may be false?

Tomas Hajek[a,b,*], Miloslav Kopecek[b], Martin Alda[a,b], Rudolf Uher[a], Cyril Höschl[b]

[a]*Department of Psychiatry, Dalhousie University, Halifax, Nova Scotia, Canada*
[b]*Prague Psychiatric Center, Department of Psychiatry and Medical Psychology, 3rd Faculty of Medicine, Charles University, Prague, Czech Republic*

Received 30 November 2012; received in revised form 15 January 2013; accepted 18 January 2013

KEYWORDS
Meta-analyses;
False negative;
Statistical heterogeneity

Abstract
Results of meta-analyses are regarded as the highest level of evidence. A statistically non-significant effect size from a meta-analysis is typically considered true negative even in the presence of a statistically significant signal in individual studies, presumed to be false positive. Here we provide examples from neuroimaging, genetics and psychopharmacology of why meta-analyses may frequently yield false negative results from true positive findings. This may happen in situations when individual studies report findings in opposing directions, the sum of which yields a non-significant overall effect size. Such non-significant meta-analyses, which show statistical heterogeneity and include studies with opposing effect sizes do not provide an accurate estimate of the overall effect and may have lower heuristic value than individual studies. Over reliance on such meta-analyses may falsely identify certain potentially fruitful research avenues as blind alleys.

1. Introduction

Meta-analytical tools are becoming increasingly popular in a growing range of disciplines. Results of meta-analyses are regarded as the highest level of evidence. A statistically non-significant effect size from a meta-analysis is typically considered true negative even in the presence of a statistically significant signal in individual studies, presumed to be false positive (Kirsch et al., 2008; Ioannidis, 2011). Here we provide examples of why meta-analyses may frequently yield false negative results from true positive findings. This may happen in situations when individual studies report findings in opposing directions, the sum of which yields a non-significant overall effect size.

2. False negative meta-analyses in neuroimaging

Eight previous meta-analyses have reported comparable hippocampal volumes in patients with bipolar disorders

*Corresponding author at: Department of Psychiatry, Dalhousie University, QEII HSC, A.J. Lane Bldg., Room 3093, 5909 Veteran's Memorial Lane, Halifax, NS, Canada B3H 2E2. Tel.: +1 902 473 8299; fax: +1 902 473 1583.
E-mail address: tomas.hajek@dal.ca (T. Hajek).

0924-977X/$ - see front matter © 2013 Elsevier B.V. and ECNP. All rights reserved.
http://dx.doi.org/10.1016/j.euroneuro.2013.01.004

(BD) relative to controls (Hajek et al., 2012b). The absence of hippocampal volume changes in BD became the dominant view of the field. Alternatively, there could be clinical or treatment related variables, which exert opposing effects on hippocampal volumes and cancel each other when analyzed jointly. Hippocampal volumes are negatively associated with duration of BD, but positively associated with Li treatment (Hajek et al., 2012a). In a meta-analysis separating patients according to Li exposure, we found that BD subjects who were currently not treated with Li had significantly smaller bilateral hippocampal volumes than healthy controls, who had significantly smaller hippocampal volumes than Li-treated BD subjects (Hajek et al., 2012b). In those meta-analyses where subjects with and without current Li treatment were analyzed jointly, the larger hippocampal volumes relative to controls in the Li-treated subjects could have masked the smaller hippocampal volumes relative to controls in the non-Li subgroups, thus yielding false negative summary effect size.

Similarly, meta-analyses have reported comparable amygdala volumes in patients with BD and controls (McDonald et al., 2004; Kempton et al., 2008). A meta-regression revealed that the mean amygdala volume difference between the BD and control subjects was significantly associated with age (Hajek et al., 2009; Usher et al., 2010). Whereas BD participants older than 20 years of age had a trend for larger amygdala volumes, BD participants younger than 20 years of age showed significantly smaller amygdala volumes than controls (Hajek et al., 2009). The interpretation of this finding is not clear. It may reflect increased likelihood of Li exposure in older subjects. Alternatively and perhaps more likely this may pertain to neurobiological differences between very early onset/pediatric BD and more classical forms of the illness or to the high rates of neurodevelopmental comorbid conditions associated with smaller amygdala volumes in studies of children/adolescent BD subjects (Hajek et al., 2009).

3. Mirror changes in genetics

This canceling of opposing effects may happen whenever mirror changes exist, which may be often. For example in genetics, environmental factors may reverse the effects of an allele (gene environment (GxE) interactions). Carriers of less active version of the monoamine oxidase A gene proved most antisocial when they experienced maltreatment in childhood, but scored lowest in antisocial behavior when not exposed to childhood maltreatment (Belsky et al., 2009). Similar crossed GxE interaction can be found for serotonin-transporter-linked polymorphic region, stressful life events and probability of suicide or major depression (Belsky et al., 2009). Interestingly, genes with documented GxE interaction for psychiatric disorders tend to show no main effect, even though typically larger samples are available for the direct association studies.

4. Opposing effects in psychopharmacology

The issues outlined above may be less problematic for systematic reviews of interventions based on randomized controlled trials. This is in part because methods to assess the quality of randomized trials are more developed than methods to assess the quality of other study designs. However, there still continues to be a marked methodological heterogeneity between randomized controlled studies (Safer, 2002) and the effect size and direction of changes may depend on the specifics of the research design (Geddes et al., 2000; Woodward et al., 2007; Safer, 2002). Opposing effects of the same treatment have frequently been documented in psychopharmacology. Antidepressants may show both antidepressant or depressogenic effects (Sharma, 2001; Ruzickova and Alda, 2002; Fava, 2003), antipsychotics may improve or worsen cognitive functions (MacQueen and Young, 2003; Woodward et al., 2007), psychostimulants may stimulate or sedate (Kratochvil et al., 2011).

5. Discussion

Non-significant results of meta-analyses, which also show high statistical heterogeneity are always suspect of being false negative, especially if individual included studies reported findings in opposite directions. Such a meta-analysis should be followed with attempts to investigate the sources of the heterogeneity by either using subgroup analyses for dichotomous moderators (Hajek et al., 2009, 2012b), or by utilizing meta-regression for continuous ones (Hajek et al., 2009; Usher et al., 2010). These methods are easy to implement. What is difficult is the selection of moderators. Considering the wealth of replicated data showing neuroprotective effects of Li, subdividing patients based on exposure to Li was a relatively clear and a priori choice in the above-mentioned example (Hajek et al., 2012b). In many other situations, such as the amygdala volumes (Hajek et al., 2009), the moderating variables may not be as obvious. Post-hoc testing of multiple moderators to explain differences between studies may not be optimal or may not succeed in reducing the statistical heterogeneity. Non-significant meta-analyses, which even after a careful investigation of moderators continue to show high levels of statistical heterogeneity, need to be considered very preliminary and of lower heuristic value than individual studies. Such meta-analyses do not represent the highest level of evidence and should not be utilized as a justification to abandon a specific research direction.

To summarize, in all of the above-mentioned situations, a non-significant effect size from a meta-analysis would be misleading and would not provide a good estimate of the real effect. Consequently, the over reliance on meta-analyses may falsely identify certain potentially fruitful research avenues as blind alleys.

Role of funding source

This study was supported by grants from the Canadian Institutes of Health Research (Grants 103703 and 106469), Nova Scotia Health Research Foundation and Dalhousie Clinical Research Scholarship to Dr. Hajek. The granting agencies had no further role in study design; in the collection, analysis and interpretation of data; in the writing of the report; and in the decision to submit the paper for publication.

Contributors

Dr. Hajek wrote the first draft of the manuscript. All authors contributed to and have approved the final manuscript.

Conflict of interest

None of the authors has any conflicts of interest to disclose.

References

Belsky, J., Jonassaint, C., Pluess, M., Stanton, M., Brummett, B., Williams, R., 2009. Vulnerability genes or plasticity genes? Mol. Psychiatry 14, 746–754.

Fava, G.A., 2003. Can long-term treatment with antidepressant drugs worsen the course of depression? J. Clin. Psychiatry 64, 123–133.

Geddes, J., Freemantle, N., Harrison, P., Bebbington, P., 2000. Atypical antipsychotics in the treatment of schizophrenia: systematic overview and meta-regression analysis. Br. Med. J. 321, 1371–1376.

Hajek, T., Cullis, J., Novak, T., Kopecek, M., Hoschl, C., Blagdon, R., O'donovan, C., Bauer, M., Young, L.T., MacQueen, G., Alda, M., 2012a. Hippocampal volumes in bipolar disorders: opposing effects of illness burden and lithium treatment. Bipolar Disord. 14, 261–270.

Hajek, T., Kopecek, M., Hoschl, C., Alda, M., 2012b. Smaller hippocampal volumes in patients with bipolar disorder are masked by exposure to lithium: a meta-analysis. J. Psychiatry Neurosci. 37, 110143.

Hajek, T., Kopecek, M., Kozeny, J., Gunde, E., Alda, M., Hoschl, C., 2009. Amygdala volumes in mood disorders—meta-analysis of magnetic resonance volumetry studies. J. Affect. Disord. 115, 395–410.

Ioannidis, J.P., 2011. Excess significance bias in the literature on brain volume abnormalities. Arch. Gen. Psychiatry 68, 773–780.

Kempton, M.J., Geddes, J.R., Ettinger, U., Williams, S.C., Grasby, P.M., 2008. Meta-analysis, database, and meta-regression of 98 structural imaging studies in bipolar disorder. Arch. Gen. Psychiatry 65, 1017–1032.

Kirsch, I., Deacon, B.J., Huedo-Medina, T.B., Scoboria, A., Moore, T.J., Johnson, B.T., 2008. Initial severity and antidepressant benefits: a meta-analysis of data submitted to the Food and Drug Administration. PLoS Med. 5, e45.

Kratochvil, C.J., Vaughan, B.S., Stoner, J.A., Daughton, J.M., Lubberstedt, B.D., Murray, D.W., Chrisman, A.K., Faircloth, M.A., Itchon-Ramos, N.B., Kollins, S.H., Maayan, L.A., Greenhill, L.L., Kotler, L.A., Fried, J., March, J.S., 2011. A double-blind, placebo-controlled study of atomoxetine in young children with ADHD. Pediatrics 127, e862–e868.

MacQueen, G., Young, T., 2003. Cognitive effects of atypical antipsychotics: focus on bipolar spectrum disorders. Bipolar Disord. 5 (Suppl. 2), 53–61.

McDonald, C., Zanelli, J., Rabe-Hesketh, S., Ellison-Wright, I., Sham, P., Kalidindi, S., Murray, R.M., Kennedy, N., 2004. Meta-analysis of magnetic resonance imaging brain morphometry studies in bipolar disorder. Biol. Psychiatry 56, 411–417.

Ruzickova, M., Alda, M., 2002. Antidepressants in bipolar depression: when less is more. J. Psychiatry Neurosci. 27, 312.

Safer, D.J., 2002. Design and reporting modifications in industry-sponsored comparative psychopharmacology trials. J. Nerv. Ment. Dis. 190, 583–592.

Sharma, V., 2001. Loss of response to antidepressants and subsequent refractoriness: diagnostic issues in a retrospective case series. J. Affect. Disord. 64, 99–106.

Usher, J., Leucht, S., Falkai, P., Scherk, H., 2010. Correlation between amygdala volume and age in bipolar disorder—a systematic review and meta-analysis of structural MRI studies. Psychiatry Res. 182, 1–8.

Woodward, N.D., Purdon, S.E., Meltzer, H.Y., Zald, D.H., 2007. A meta-analysis of cognitive change with haloperidol in clinical trials of atypical antipsychotics: dose effects and comparison to practice effects. Schizophr. Res. 89, 211–224.