Letter to the Editor: Claims about the effects of botulinum toxin on depression should raise some eyebrows.

Nicholas A. Coles[a] and Jeff T. Larsen[b]

[a] Center for the Study of Language and Information, Stanford University, Palo Alto, USA; ncoles@stanford.edu

[b] Department of Psychology, University of Tennessee, Knoxville, USA; jeff.larsen@utk.edu

Corresponding author: Nicholas A. Coles; ncoles@stanford.edu

Formatted copy of article is available in the *Journal of Psychiatric Research*

https://doi.org/10.1016/j.jpsychires.2021.05.021

Recently, Schulze and colleagues (2021) and we (Coles, Larsen, Kuribayashi, & Kuelz, 2019) published two separate meta-analyses examining whether glabellar-region botulinum toxin injections can decrease depression. Both meta-analysis teams reviewed similar studies; discussed similar mechanisms-of-action; observed unexpectedly large effect sizes; observed asymmetry in funnel plot distributions; and acknowledged that it is difficult to blind participants. Yet, our two teams came to starkly different conclusions. Whereas Schulze and colleagues concluded that the treatment reaches rigorous "1a level of evidence" standards (p. 338), we concluded the opposite: that the claim is "not yet well substantiated by a credible balance of evidence" (p. 11).

How is it that our two meta-analysis teams made similar observations but formed opposite conclusions? We explore this issue further in this Letter to the Editor.

Surprisingly Large Effect Sizes

When evaluating the evidence for this new treatment, we can use prior knowledge of the underlying theory and other depression interventions to establish a range of plausible effect sizes. If we base our expectations on what both meta-analysis teams state is the underlying theory—the theorized effects of facial feedback on emotion (Coles et al., 2021)—previous meta-analyses suggest that we should expect an extremely small effect ($d = 0.17$; Coles, Larsen, & Lench, 2019). If we base our expectations on prior knowledge of other pharmacological depression interventions, previous reviews indicate that we should expect a medium-sized effect ($d = 0.42$; Turner, Matthews, Linardatos, Tell, & Rosenthal, 2008). Based on these expectations, both meta-analysis teams acknowledge that the reported effect of botulinum toxin on depression is surprisingly large ($d > .80$). These results seem to suggest that the treatment is over 4 times the size of typical facial feedback effects and 2 times the size of established antidepressant

treatments. Such a large effect would revolutionize the treatment of depression and the study of facial feedback theory—if real.

Unfortunately, effect sizes that drastically exceed theory- and practice-based expectations are often more cause for concern than celebration (Gelman & Carlin, 2014; Hilgard, 2021). Such results indicate that there may be a problem with the research design, the proposed mechanism-of-action, and/or the analysis or reporting of the data. Both of our meta-analysis teams acknowledge that issues with participant blinding and placebo effects may account for the surprisingly large effect size estimate. Yet, our two meta-analysis teams come to different conclusions. Schulze and colleague argued that we addressed issues with participant blinding in our own meta-analysis. We, however, actually concluded the opposite: that these issues potentially undermine claims about the effects of botulinum toxin on depression.

Alternative Mechanisms-of-Action

In addition to research design issues, misspecified mechanisms-of-action can account for surprisingly large effect size estimates. Both of our meta-analysis teams agree that the hypothesized effect of botulinum toxin on depression is derived from facial feedback theory. However, both of our teams also discuss several other plausible mechanisms, including improvements in appearance, quality of life, and social treatment.

Both of our teams review preliminary evidence that is inconsistent with the appearance mechanism: a pooled analysis that found that changes in frown line intensity did not significantly predict botulinum toxin recipients' improvements in depression (Reichenberg et al., 2016). Schulze and colleagues additionally argue that preliminary evidence is inconsistent with the quality of life and social treatment explanations. Specifically, Schulze and colleagues argue that

evidence from a small Master's thesis by Sharif (2013) suggests that botulinum toxin users "indicated feeling more rejected by others" (p. 334). This conclusion is puzzling, however, considering that Sharif did *not* find that perceived social rejection significantly differed in the botulinum toxin vs. placebo groups.

We believe that ongoing uncertainty about mechanism of action has important social, ethical, and practical implications. Imagine that botulinum toxin does indeed decrease depression, but only because it improves subjective appearance. What social forces are creating this link between appearance and depression? Should doctors reinforce these social forces by providing cosmetically-mediated depression treatments (Chatterjee, 2007)? Should insurance companies pay for these types of treatments? Given the implications of these highly plausible alternative explanations, we believe that rigorous examinations of these accounts are needed.

Funnel Plot Asymmetry

Both of our meta-analysis teams observed asymmetry in the effect size funnel plot distributions, which can have benign (e.g., statistical artifacts) or malignant causes (e.g., publication bias; Sterne et al., 2011). We warned that the funnel plot asymmetry is concerning because (1) 51% of the relevant effect sizes were missing from the published literature, and (2) 96% of effect sizes came from investigators with conflicts of interest. Schulze and colleagues, on the other hand, suggested that the funnel plot asymmetry is a statistical artifact driven by "variability in the behavior of the placebo groups" (p. 338). Schulze and colleagues' explanation, however, is puzzling because variability in the placebo group response—by itself—should not lead to funnel plot asymmetry. Instead, funnel plot asymmetry should only emerge if the placebo groups systematically had more exaggerated responses in the larger vs. smaller studies. Schulze

and colleagues provide little evidence to support such an assertion. Thus, we believe that the observed funnel plot asymmetry continues to be concerning.

Concluding Remarks

Our two meta-analysis teams most strikingly disagree about whether botulinum toxin should be promoted as a depression treatment considering concerns about unusually large effects, unaddressed blinding issues, unclear mechanisms-of-action, large proportions of missing data, preponderances of conflicts of interest, and funnel plot asymmetry.

If this treatment is effective and the proposed mechanism-of-action is correct, our relative pessimism may deprive patients of a potentially useful depression treatment. If this treatment is not effective or the mechanism-of-action is misspecified, Schulze and colleagues' relative optimism may cause patients to incur unnecessary financial costs; forgo more established depression treatments; unnecessarily expose themselves to (admittedly uncommon) adverse reactions; and/or unnecessarily seek cosmetic botulinum toxin treatments due to its purported effects on mood. We believe that a more careful consideration of the quality of the evidence, potential costs, and potential benefits is necessary before promoting botulinum toxin as an off-label treatment for depression.

References

Chatterjee, A. (2007). Cosmetic neurology and cosmetic surgery: Parallels, predictions, and challenges. *Cambridge Quarterly of Healthcare Ethics*, *16*(2), 129–137.

Coles, N. A., Larsen, J. T., Kuribayashi, J., & Kuelz, A. (2019). Does blocking facial feedback via botulinum toxin injections decrease depression? A critical review and meta-analysis. *Emotion Review*, *11*(4), 294–309.

Coles, N. A., Larsen, J. T., & Lench, H. C. (2019). A meta-analysis of the facial feedback literature: Effects of facial feedback on emotional experience are small and variable. *Psychological Bulletin*, *145*(6), 610–651.

Coles, N. A., March, D. S., Marmolejo-Ramos, F., Banaruee, H., Butcher, N., Cavallet, M., … Gorbunova, E. (2021). A multi-lab test of the facial feedback hypothesis by The Many Smiles Collaboration. Retrieved from https://psyarxiv.com/cvpuw

Gelman, A., & Carlin, J. (2014). Beyond power calculations: Assessing Type S (Sign) and Type M (Magnitude) errors. *Perspectives on Psychological Science*, *9*(6), 641–651.

Hilgard, J. (2021). Maximal positive controls: A method for estimating the largest plausible effect size. *Journal of Experimental Social Psychology*, *93*, 1–10.

Reichenberg, J. S., Hauptman, A. J., Robertson, H. T., Finzi, E., Kruger, T. H. C., Wollmer, M. A., & Magid, M. (2016). Botulinum toxin for depression: Does patient appearance matter? *Journal of the American Academy of Dermatology*, *74*(1), 171–173.

Schulze, J., Neumann, I., Magid, M., Finzi, E., Sinke, C., Wollmer, M. A., & Krüger, T. H. C. (2021). Botulinum toxin for the management of depression: an updated review of the

evidence and meta-analysis. *Journal of Psychiatric Research*, *135*, 332–340.

Sharif, V. (2013). *Does Botox buffer the negative effects of social rejection?: A test of the facial feedback hypothesis.* (Master's Thesis, University of Kentucky, Kentucky, USA). Retrieved from https://uknowledge.uky.edu/psychology_etds/25/

Sterne, J. A. C., Sutton, A. J., Ioannidis, J. P. A., Terrin, N., Jones, D. R., Lau, J., … Higgins, J. P. T. (2011). Recommendations for examining and interpreting funnel plot asymmetry in meta-analyses of randomised controlled trials. *BMJ*, *342*, 1–8.

Turner, E. H., Matthews, A. M., Linardatos, E., Tell, R. A., & Rosenthal, R. (2008). Selective publication of antidepressant trials and its influence on apparent efficacy. *New England Journal of Medicine*, *358*(3), 252–260.