Authors' Response to a Critique by Jauhar and Hayes of 'a Systematic Review into the Incidence, Severity and Duration of Antidepressant Withdrawal Effects: Are Guideline Evidence-based?'

James Davies, John Read

PII: S0306-4603(19)30092-9
DOI: https://doi.org/10.1016/j.addbeh.2019.01.026
Reference: AB 5881

To appear in: *Addictive Behaviors*

Please cite this article as: J. Davies and J. Read, Authors' Response to a Critique by Jauhar and Hayes of 'a Systematic Review into the Incidence, Severity and Duration of Antidepressant Withdrawal Effects: Are Guideline Evidence-based?', Addictive Behaviors, https://doi.org/10.1016/j.addbeh.2019.01.026

Authors' Response to a Critique by Jauhar and Hayes of 'A Systematic Review into the Incidence, Severity and Duration of Antidepressant Withdrawal Effects: Are Guideline Evidence-Based?'

Commentary

James Davies[a,b,*] james.davies@stx.oxon.org and John Read[c,d]

[a]University of Roehampton, London, United Kingdom
[b]All-Party Parliamentary Group for Prescribed Drug Dependence, United Kingdom
[c]University of East London, London, United Kingdom
[d]International Institute for Psychiatric Drug Withdrawal, Sweden

[*]Corresponding author.

Hayes and Jauhar's first commentary on our review (Davies and Read 2018a) took the form of a blog (Hayes and Jauhar 2018). Many criticisms raised in that blog were later shown to be in error (Davies and Read 2018b), which presumably explains why they were not incorporated into their current journal commentary. Nonetheless, we will use this space to respond to both the blog and the journal commentary. Our response will clearly evidence that Hayes and Jauhar commit so many serious mistakes and misrepresentations (and/or misunderstandings) in both their blog and their commentary, that their critiques are imbalanced, imprudent and inaccurate.

Search strategy and 'biased' selection

In both their blog and commentary Hayes and Jauhar (2018) allege that it is 'highly likely' that our search strategy did not find all relevant studies. We are confident that our strategy, which not only included *MEDLINE/PubMed*, *PsycINFO* and *Google Scholar*, but previous reviews and the bibliographies of 20 relevant papers, is at least as thorough as most systematic reviews with regards to identifying published studies. We have also subsequently searched for relevant unpublished theses, dissertations and conference proceedings, in *ProQuest* and *OpenGrey*, and found none.

To support their contention that our search strategy excluded important studies, Hayes and Jauhar pronounce, in blog and commentary, that we failed to include the following five RCTs (i.e. Baldwin et a. 2004a & 2004b; Lader et al. 2004; Montgomery et al. 2003 & 2004). They alleged that these five trials, while primarily focusing on the effectiveness of antidepressants, also contained data on the 'incidence' of withdrawal – that is, on how common withdrawal actually is. They contest that if we had included these studies our incidence rates would have been lower, perhaps by around 10%. As this is a core argument of their critique, we must inspect it carefully.

Had Hayes and Jauhar read these so-called studies, even cursorily, they would have stopped in their tracks, as we did. Firstly, all five studies were written (either entirely or in part) by employees of Lundbeck Pharmaceuticals (who, in these studies, were researching their own drugs). Secondly, three of the five were not published as studies at all (so we can't even assess their methodology). Rather they were published as short (300 word) research 'supplements' (i.e. industry-funded study-summaries that some journals will publish in return for an industry fee). Needless to say, the obvious conflicts of interests these supplements involve (Lundh et al. 2010) as well as the serious challenges they pose to anyone wanting to assess their methods properly (supplements don't provide enough detail for that), are just two among numerous ethical and scientific reasons why many credible journals, such as *Lancet*, now refuse to publish them (Lancet 2010).

Finally, and most damningly for Hayes and Sameer, *none of the five so-called studies contain the incidence data they quote in their critique*. To repeat, these five studies do not contain the very data that Hayes and Jauhar alleged we overlooked.

While this, of course, explains why we did not include these studies in our systematic review, it does not explain why Hayes and Jauhar claimed the data were there. We can only surmise that they did not actually check these five studies. Rather, they took the shortcut of quoting a Lundbeck-funded article, published three years later (by Baldwin et al. 2007), which somehow 'cites' data from these original five studies that were never included in them.

By basing their arguments on such spurious foundations, Hayes and Jauhar not only demonstrate a concerning lack of caution, but also invalidate many of their core conclusions. For instance, they invalidate their argument that withdrawal incidence rates of 10-12% are not outliers. They also invalidate their re-analysis of overall incidence rates and their pooled figure for incidence from the RCTs (Table 1 of their commentary).

Inclusion and exclusion criteria

Hayes and Jauhar accurately point out (in both blog and commentary) that we did not register our inclusion criteria in advance. PRISMA does not require such criteria to be preregistered (see PRISMA guideline item #5), while we clearly state our eligibility criteria (which complies with PRISMA guideline item # 6. Furthermore, we address PRISMA items #7-#10 in the methods section). We surpassed most systematic reviews on depression treatments, as only around 30% (Chung et al., 2018) include full lists of included and excluded studies, as we do.

Hayes and Jauhar then suggest (in their blog) that we should have identified the length of follow-up, length of antidepressant exposure, and drug-company funding as reasons for exclusion before undertaking the search and data extraction. Firstly, we went beyond both what PRISMA requests and the procedure of previous systematic reviews in this area, by explicitly stating in each relevant section a rationale for excluding each of the studies omitted. Because we did so, any diligent and impartial reader will clearly see that none of the three variables that Hayes and Jauhar raise were the sole reason for excluding any individual study at all (hence our not identifying them as sole reasons for exclusion before our search and data extraction). For example, as our tables make clear, six of the *included* 24 studies were identified as drug-company funded. Furthermore, the five drug-company studies excluded from our estimate of incidence, while indeed reporting artificially short durations, were *excluded* on the quite obvious ground that they failed to report incidence rates. This was all very plainly stated, even though Hayes and Jauhar imply otherwise.

Hayes and Jauhar also take issue (in their blog) with our excluding two studies from our estimates of incidence “because they assessed only 9 withdrawal symptoms”. This

is, once again, misleading. We did not exclude them on this basis alone but, as explicitly stated, for a variety of methodological considerations. Both studies were chart-reviews, which are notoriously weak owing to reliance on practitioners being aware of, and recording, withdrawal reactions; while one study, oddly enough, excluded any withdrawal reactions commencing three days after discontinuation.

We note that Hayes and Jauhar only find reasons to challenge the exclusion of studies with relatively low incidence rates, but do not find fault with, or even acknowledge, our exclusion of a study with a 97% incidence rate.

Estimates of severity

In the following statement Hayes and Jauhar make yet another error: "When addressing severity of withdrawal the authors only consider survey data. The reason given is that many of the randomised trials and naturalistic studies are short-term, or at risk of potential bias owing to conflicts of interest".

The confidence with which they assert this falsehood is, in our view, concerning. Once again, had Hayes and Jauhar carefully read these primary sources, they would have seen that our reasons for only considering survey data when assessing severity were entirely different from the ones they allege. We only considered survey data on severity, *because the RCTs did not provide any data on the severity of withdrawal effects*.

So why do Hayes and Jauhar suggest otherwise? Firstly, they may wrongly believe that the data on severity of *depression* pre- and post-discontinuation in Rosenbaum et al. (1998) is data about severity of withdrawal effects, which it is not. Secondly, the one RCT that did provide severity data on withdrawal (Sir et al., 2005) was, as clearly stipulated in our review, excluded for two reasons: because it only covered eight weeks treatment (which would significantly lower severity rates), and because it was clinician-rated rather than self-report. Hayes and Jauhar don't acknowledge that we also stated that even if we had included this outlier, the weighted average of people who described their withdrawal effects as severe would have reduced only slightly - from 45.7% to 43.5%.

Furthermore, in their blog, Hayes and Jauhar misrepresented our 45.7% weighted average by wrongly stating that we had concluded that: 'The severity of these symptoms is severe in over half of cases'. We should not have to point out that 45.7% does not constitute 'over-half'.

Assessment of study quality and concerns about surveys

Hayes and Jauhar state in both commentary and blog that 'This review does not attempt a traditional assessment of bias in the studies they include'. In fact, we described the methodology of every one of the 24 studies in text and tables, so that readers can assess for themselves their quality, including any sample biases.

Hayes and Jauhar argue that it is questionable to combine data from randomised controlled trials and naturalistic studies with survey data (which makes us ask their own 're-analysis' did precisely this – Figure 1). As RCTs and naturalistic studies are regularly covered in systematic reviews (Guyatt et al. 2008; Egger et al. 2001), they obviously object to our inclusion of surveys. In making this objection, however, Hayes and Jauhar are simply confusing a methodological preference for a methodological law. There is no law prohibiting the inclusion of experiential survey data in a systematic review. Furthermore, our 'Limitations' section clearly acknowledges both the potential minimising bias of the RCTs because of their artificially short treatment and follow-up durations (about which Hayes and Jauhar express no concern), and the potential maximising bias of the surveys because they may attract a disproportionate number of people unhappy with their drugs (about which Hayes and Jauhar express grave concern). We also pointed out, however, that surveys can be prone to bias either way – e.g. the largest survey contained an unusually high proportion (82%) of people who thought the drugs had helped them (Read et al. 2014), so it is feasible, in this case, that the sample bias may have been towards people with a generally *positive* attitude about antidepressants, and therefore the study *underestimated* adverse effects such as withdrawal. While the RCTs had extremely artificial samples and conditions (and small numbers) the large online surveys, while not necessarily representative of all users (like the RCTs), something we do acknowledge, they do represent the *real-life* experiences of several thousand people with a range of treatment durations (from weeks to years) and various speeds of withdrawal. Psychiatry has too often been guilty of devaluing the importance of

experiential knowledge in its evaluation of interventions, which has in turn undermined our capacity to intervene effectively.

In any case, how different would our incidence estimate have been had we omitted surveys from our analysis? The three types of studies, when grouped, did not differ greatly in terms of withdrawal incidence. The weighted averages are as follows:

- The three surveys – 57.1% (1790/3137),
- The five naturalistic studies – 52.5% (127/242)
- The six RCTs – 50.7% (341/673)

Reaching similar findings from different methodologies is typically seen to strengthen confidence in an overall, combined estimate, and not, as Hayes and Jauhar imply, weaken confidence. In fact, findings from the three methodology types demonstrate that it is broadly safe to conclude that at least half of people suffer withdrawal symptoms when trying to come off antidepressants.

Outcome measures

Hayes and Jauhar (in blog and commentary) rightly state that in three of the incidence studies that we reviewed some withdrawal symptoms (as identified by the DESS) were also present in some of those continuing to take antidepressants (we will explain to them in the 'placebo' section below the quite obvious reason why some studies may find this). However, Hayes and Jauhar wrongly claim that in the Zajecka (1998) study withdrawal incidence is 'higher' in those continuing to take antidepressants compared to those stopping. The difference between the two overall rates was clearly stated in the study as not significantly different. Furthermore, four specific withdrawal effects (dizziness, dysmenorrhea, rhinitis and somnolence) *were* significantly more common in participants who had come off the drugs.

Hayes and Jauhar are also correct that the studies used different cut-off points on the Discontinuation-Emergent Signs and Symptoms Scale (DESS). We reported those differences faithfully in Table 1. Furthermore, we share their concern that academic psychiatry has failed to focus sufficiently on changes in withdrawal symptoms rather

than just presence of symptoms. We disagree, however, with their speculation (backed by no references) that such studies, past or future, lead to 'very small increases in the DESS' and that the pharmacological contribution to withdrawal is therefore 'mild and transient.'

Placebo

Hayes and Jauhar, in their commentary, are concerned that some studies find withdrawal symptoms in people who have been on placebos. While this was not the case for the majority of studies we reviewed, it was the case for one study they select as an example - Montgomery et al. (2005) – about which they state: the 'DESS score was higher during placebo treatment than in active treatment'. Hayes and Jauhar then proceed to conclude, oddly enough, that such findings somehow 'illustrate precisely what a nocebo effect is' and that by allegedly missing this we 'misunderstand simple principles' about RCTs.

This curious and erroneous criticism appears to be rooted in their confusion about the study. So, to clarify, the Montgomery study for them: after 12 weeks of open-label treatment, people were randomly and blindly assigned to either placebo or to continuation on the drug. As those randomised to placebo therefore *stopped* the active drug, *this explains why many randomised to placebo suffered withdrawal effects* (and so clearly such withdrawal does not 'illustrate' nocebo). But what can explain withdrawal effects being found in 9% of those who remained on the antidepressant? If Hayes and Jauhar imply nocebo effects can explain this, once again they are in error. As a careful reading of the study shows, it does not control for the fact that experiences classed by DESS as 'withdrawal' may actually be 'treatment-emergent adverse events' (as classified by TEAE). As both DESS and TEAE overlap in some key symptoms they classify (e.g. nausea, sweating, anxiety, dizziness, insomnia) it is possible that the 9% of people labelled by DESS as experiencing 'withdrawal' were rather experiencing 'treatment-emergent adverse events'. This explains the oddity of so-called withdrawal occurring in the treatment group – i.e. such 'withdrawal' may not be withdrawal at all but an artefact of misclassifying it as such. It is strange that Hayes and Jauhar could overlook this quite basic point when putting together their commentary.

Combining data from different types of antidepressants

Hayes and Jauhar (in blog and commentary) argue that different antidepressant medications have different withdrawal effects and that it was 'a major conceptual failure to not address it in the review'. Firstly, in our systematic review we explicitly acknowledge that "differing half-lives affect timing of withdrawal onset", so they are telling us nothing we don't already declare. Secondly, our Table 1 (on incidence) actually lists rates for all the individual drugs provided by all the studies. Furthermore, providing global estimates was both necessary and appropriate given the central aim of our review. Our aim was to assess whether NICE guidelines (2009) on antidepressant withdrawal were evidence based, not to guide clinicians in what specific drugs to prescribe nor to illuminate the particularities of different pharmacokinetic properties. Here Hayes and Jauhar commit the all-too-common fallacy of criticising a study for not doing what it never set out to do.

'Minor errors'

We thank Hayes and Jauhar for identifying three very 'minor errors' in presentation, which do not impact our estimates, and which have been corrected.

We are concerned, however, that they purport to identify two further 'errors', which clearly represent errors on their part not ours. Firstly, their assertion that the 'total number experiencing withdrawal in the study by Sir and colleagues is 83 rather than 110' is wrong. They reached this figure by unjustifiably removing all withdrawal symptoms rated 'minimal', while failing to inform readers that they did so.

Their second error concerns their suggestion that we misrepresented the incidence rate of another study (Montgomery et al. 2005) by presenting the incidence of withdrawal following escitalopram treatment as '27%, when it is 16%.' Here Hayes and Jauhar again mislead the reader. The 27% rate we reported was at one week and the 16% they reported was at two weeks. Given we were calculating for incidence, it was absolutely correct for us to use the 27% figure in our calculations.

These are unfortunate errors for Hayes and Jauhar to make, suggesting they did not subject their commentary to the rigorous level of quality checking that a serious readership should expect.

Conclusion

For the many reasons stated above Hayes and Jauhar's commentary turns out to be both inaccurate and misleading. In some cases, the critiques they offer are based on stark misrepresentations of study findings and/or the misreading or non-reading of primary sources.

Furthermore, we also note that every single criticism, error and/or misrepresentation that Hayes and Jauhar made, resulted in the minimisation (i.e. the downplaying) of both the incidence and severity of antidepressant withdrawal effects. Despite their commentary being biased in this direction, their own re-analysis, which inexcusably includes studies that *provide no data on incidence at all*, remarkably produces an overall withdrawal incidence rate of 44% (Figure 1), which (despite underestimating the actual rate) would still represent 3.2 million adults in England alone.

We accept that our overall estimates of 56% incidence, with 46% of those being severe, are indeed only estimates. Yet, even if the actual incidence is towards the lower end of the 50% to 57% range, when grouping study types, this would still constitute at least half of all antidepressant users (at least 3.6 million adults in England alone). It is crucial that amid the complexities of academic disagreement we do not lose sight of the scale of the problem that our systematic review helps to expose.

In the light of this, it is also interesting to note the absence of any acknowledgment by Hayes and Jauhar that we are discussing a public health issue involving millions of people worldwide. They also fail to comment on the primary finding of the review, namely that national guidelines in the U.S.A. and the U.K. significantly misjudge the true extent of the problem.

In the spirit of seeking some common ground between us, we can agree, however, with one statement Hayes and Jauhar made in their original blog:

> 'It reflects negatively on the whole of the field of psychiatry that there is not better, clearer evidence from high quality studies on the incidence, severity and duration of any symptoms related to antidepressant cessation' (Hayes and Jauhar 2018).

Given that 16% of the English adult population was prescribed an antidepressant last year alone (7.3 million people), this professional oversight, and its significance, is hard to excuse.

While better research is indeed desirable (with respect to a diversity of issues pertaining to withdrawal), the millions of people experiencing withdrawal effects cannot wait for psychiatry to determine whether they represent 50% or 57% of those withdrawing, or what will be the best methodologies to more precisely assess that. They need accurate information and proper support now. And the millions more who will consider starting antidepressants in the coming years are entitled, unlike those who have gone before, to receive accurate information about all the adverse effects including the difficulty they are very likely to encounter when they try to stop; difficulties that in many cases will be protracted and severe.

A crucial step forward will be for government bodies, and professional organisations, to update their guidelines so as to render them evidence-based, and thereby of maximum benefit to the public.

References

Baldwin, D.S., Hindmarch. I., Huusom A.K.T., Cooper, J. (2004a). Escitalopram and paroxetine in the short and long-term treatment of major depressive disorder (MDD*). International Journal of Neuropsychopharmacology,* 7 (Suppl. 2), S168–S169.

Baldwin, D.S., Huusom, A.K.T., Mæhlum, E. (2004b). Escitalopram and paroxetine compared to placebo in the treatment of generalised anxiety disorder (GAD). *European Neuropsychopharmacology,* 14 (Suppl. 3), S311.

Baldwin, D.S., Montgomery, S.A., Nil, R., Lader, M. (2007). Discontinuation symptoms in depression and anxiety disorders. *International Journal of Neuropsychopharmacology*, 10(1):73-84.

Chung, V.C.H., Wu, X.Y., Feng, Y., Ho, R.S.T., Wong, S.Y.S., Threapleton, D. (2018). Methodological quality of systematic reviews on treatments for depression: a cross-sectional study. *Epidemiology and Psychiatric Sciences*, 27 (26): 619-627.

Davies, J., & Read, J. (2018a). A systematic review into the incidence, severity and duration of antidepressant withdrawal effects: Are guidelines evidence-based? (PDF) *Addictive Behaviors*. Sep 4. https://doi.org/10.1016/j.addbeh.2018.08.027

Davies, J, Read, J (2018b). Antidepressant withdrawal review: authors respond in detail to Mental Elf critique. Council for Evidence-based Psychiatry. Website http://cepuk.org/2018/11/05/antidepressant-withdrawal-review-authors-respond-mental-elf-critique/ (accessed Dec 2019).

Egger, M., Davey-Smith, G., Altman, D. (eds.) (2001) *Systematic reviews in health care: meta-analysis in context.* 2nd ed. London (UK): BMJ Publishing Group.

Guyatt, G., Rennie, D., Meade, M., Cook, D. (2008) *Users' guides to the medical literature*. 2nd ed. New York (NY): McGraw Hill Medical.

Hayes, J. & Jauhar, S. (2018) Antidepressant withdrawal: reviewing the paper behind the headlines. Mental Elf. Website https://www.nationalelfservice.net/treatment/antidepressants/antidepressant-withdrawal-reviewing-the-paper-behind-the-headlines/ (accessed Dec 2018).

Lader, M., Stender, K., Burger, V., Nil, R. (2004). The efficacy and tolerability of escitalopram in 12- and 24-week treatment of social anxiety disorder: a randomised, double-blind, placebo-controlled, fixed-dose study. *Depression and Anxiety,* 19, 241–248.

The perils of journal and supplement publishing. *The Lancet* 2010; 375(9712): 347. DOI:https://doi.org/10.1016/S0140-6736(10)60147-X

Lundh, A., Barbateskovic, M., Hróbjartsson, A., Gøtzsche, P.C. (2010) Conflicts of interest at medical journals: the influence of industry-supported randomised trials on journal impact factors and revenue - cohort study. *P L o S Medicine.* 2 (1);7(10):e1000354. https://doi.org/10.1371/journal.pmed.1000354

Montgomery, S.A., Durr-Pal, N., Loft, H., Nil, R. (2003). Relapse prevention by escitalopram treatment of patients with social anxiety disorder (SAD). *European Neuropsychopharmacology,* 13 (Suppl. 4), S364.

Montgomery, S.A., Huusom, A.K.T., Bothmer, J. (2004a). A randomised study comparing escitalopram with venlafaxine XR in patients in primary care with major depressive disorder. *Neuropsychobiology,* 50, 57–64.

Montgomery, S., Nil, R., Durr-Pal, N., Loft, H., & Boulenger, J. (2005). A 24-week randomized, double- blind, placebo-controlled study of escitalopram for the prevention of generalized social anxiety disorder. *Journal of Clinical Psychiatry*, 66, 1270–1278

Read, J., Cartwright, C., & Gibson, K. (2014). Adverse emotional and interpersonal effects reported by 1,829 New Zealanders while taking antidepressants. *Psychiatry Research*, 216, 67–73.

Rosenbaum, J. F., Fava, M., Hoog, S. L., Ascroft, R. C., & Krebs, W. B. (1998). Selective serotonin reuptake inhibitor discontinuation syndrome: A randomized clinical trial. *Biological Psychiatry*, 44, 77–87.

Sir, R. F., D'Souza, S. U., George, T., Vahip, S., Hopwood, M., Martin, A. J., ... Burt, T. (2005). Randomized trial of sertraline versus venla- faxine XR in major depression: Efficacy and discontinuation symptoms. Journal of Clinical Psychiatry, 66, 1312–1320.

Zajecka, J., Fawcett, J., Amsterdam, J., Quitkin, F., Reimherr, F., Rosenbaum, J., ... Beasley, C. (1998). Safety of abrupt discontinuation of fluoxetine: A randomized, placebo-controlled study. *Journal of Clinical Psychopharmacology*, 18(3), 193–197.