Withdrawal effects confounding in clinical trials: another sign of a needed paradigm shift in psychopharmacology research

David Cohen [ID] and **Alexander Recalt** [ID]

Abstract: Randomized controlled trials' ability to produce evidence useful for people to decide whether to take, continue taking, or stop taking psychotropic drugs has been intensely critiqued, along with the trials' commercial, ideological, and regulatory contexts. This article applies the critique to the topic of withdrawal effects confounding the outcomes of relapse-prevention trials where prescribed psychotropic drugs are discontinued. Until recently, the complexity and reach of withdrawal and post-withdrawal effects were neglected by mainstream psychiatry, but not by lay users of prescribed psychotropics. This article discusses withdrawal effects as part of the pharmacology of psychotropic drugs but shaped by psychosocial factors, and possibly shaping the presentation of psychological distress generally. It outlines biases and misconceptions in assumptions, design, and reporting of general efficacy trials and findings from a recent review of 80 discontinuation trials. In theory, relapse-prevention trials are tautological and exaggerate efficacy. In publications, they pay little attention to the central feature of their design, favor abrupt or rapid discontinuations, do not attend to environmental factors, and provide insufficient data to allow re-analyses. Thus, relapse-prevention RCTs likely confound the detection of their main outcome of interest: "relapse." Using slower tapers, active placebo controls, and specific covariates in analyses would reduce the risk of withdrawal confounding, and better reporting would reduce the opaqueness of trials. The crisis in psychopharmacology is fueled partly by the disconnect between claims of therapeutic efficacy from so-called best-evidence methods despite unchanging population-level indicators of psychiatric sickness. Only by "stacking the deck" against trial sponsors' hoped-for outcomes can psychopharmacology trials regain scientific credibility.

Keywords: drug industry, outcome measurement error, placebo, randomized controlled trials, substance withdrawal syndrome

Introduction

In this article, we argue that the detection of relapse in "relapse-prevention" randomized controlled trials (RCTs) of prescribed psychotropic drugs is confounded by the withdrawal effects of these drugs, which confers on them an unjustified advantage over inert placebo. Withdrawal confounding touches at the heart of psychopharmacological efficacy: in a 2003 systematic review of 31 longer-term trials (cited more than 1000 times), which concluded that continuing antidepressant treatment halved the risk of relapse compared with placebo,[1]

the authors cautioned that because included studies "unavoidably ... necessitated" discontinuing some patients from active treatment and giving them placebos, "the risk of relapse or recurrence might be increased by a direct quasi-pharmacological response to the withdrawal of medication *per se* rather than the relapse or recurrence being solely due to the underlying disorder" (p. 660). Yet, researchers and regulators such as the United States Food and Drug Administration (FDA) have so far ignored withdrawal confounding or accepted its occurrence, and strategies to exploit

Correspondence to:
David Cohen
UCLA Luskin School of
Public Affairs, 3250 Public
Affairs Building, Los
Angeles, CA 90095-1656,
USA
cohen@luskin.ucla.edu

Alexander Recalt
Department of Social
Welfare, Luskin School
of Public Affairs,
and Department of
Epidemiology, Fielding
School of Public Health,
University of California,
Los Angeles, CA, USA

its occurrence, as normal science. While future studies could greatly reduce the risk of withdrawal confounding, it is but one sign of a broader dysfunction in the pre-market approval process of psychotropic drugs, which, in its final phase, relies on RCTs.

The drug approval process aims to determine a *prescribed* psychoactive drug's *efficacy* to treat a *mental disorder*. The italicized words mean, respectively, that the drugs have potentially dangerous uses and are expected to be used under some medical supervision, and that their therapeutic benefits and risks will be assessed formally in relation to target conditions that justify using or marketing the drugs. These formal processes differ from those where people informally decide whether and how psychoactive drugs alleviate personal distress (e.g., help them sleep after suffering a loss) and affect them more generally (e.g., alter their consciousness). Most adults have conducted such evaluations regarding some psychoactive drug(s), and their experiences strengthen assumptions about drug effects that clash with the sort of evidence RCTs are designed to produce.[2,3] The assumptions are that desirable (therapeutic) and undesirable (adverse) effects of psychotropic drugs, including withdrawal effects, are complex and variable in human beings. The complexity is definitional: substances that affect the central nervous system and alter thinking, feeling, and behaving probably vary in their effects (in degree and kind) across individuals, within individuals, and over time.

However, recent tests have not found evidence that the intended effects of antipsychotics and antidepressants – hypothesized and quantified as changes in outcome measures (numerical symptom rating scores) – vary from patient to patient.[4,5] We believe the tests are hamstrung by the outcome measures, which must narrow the possible range of experiences with psychotropic drugs in order to detect a drug effect on a presumed discrete disease entity. Yet, how to interpret the *meaning* of symptom reduction on a rating scale is far from self-evident, given the general muting of the patient's voice in RCTs.[6] While measures of "patient reported outcomes" exist, they are not yet prominently incorporated as primary outcomes in psychopharmacology RCTs, and, by relying on standardization and quantification to establish reliability and validity, may fail to fully capture the person's experience. Therefore, while the RCT might well help to determine whether a

given psychoactive drug can demonstrate statistically to a regulatory agency "a positive effect on some area of concern, some proportion of the time",[7] the RCT is ill-suited to determine whether that drug taken for long periods safely helps people overcome or palliate distress, and even less suited to determine how else the drug might affect people as they go about their lives, and to help them decide whether to continue taking or cease taking the drug.

Psychopharmacology RCTs exist within several interwoven contexts, three of which unequivocally shape the purposes of these experiments. First, biopsychiatry, especially in America, has committed itself to viewing any serious human distress that does not remit easily as an idiopathic somatic disease. Despite countless critiques and acknowledgements of failure to validate this view,[8–10] the commitment remains unwavering. Second, the pharmaceutical industry dominates the testing of psychiatric drugs by funding it directly and indirectly, while controlling most distribution channels for drug information.[8,11] Calls to shift funding of clinical trials to primarily public sources have generally fallen on deaf ears,[12] despite the existence of the third context: an ever-growing proportion of adults (about 1 in 5) and children (about 1 in 10) taking, most for the long term, psychopharmaceuticals prescribed to them by physicians.[13,14] Thousands of psychopharmacology RCTs with "positive" results have been published in medical journals. Yet, according to the former head of the world's largest government funder of mental health research, there has been, over the past decades at the population level, "no evidence of reduced morbidity or mortality from any mental illness" (p. 129)[15] – no reduced occurrence, no shortened episodes, no less frequent relapses, no increased life expectancy, no fewer suicides.

Psychotropic drug discontinuation and withdrawal effects

The word "withdrawal" has long been used to characterize symptoms that appear after stopping a prescribed centrally active or psychotropic drug and are attributed to the actions of the drug.[16] The alternative word "discontinuation" was promoted by the pharmaceutical industry to ensure that selective serotonin reuptake inhibitors (SSRIs) were not seen by physicians and the public as addictive,[17–19] but reports of contemporary RCTs use both words interchangeably; "withdrawal" is

more common in studies of benzodiazepines and stimulants, "discontinuation" in studies of antidepressants and antipsychotics.[20] To reduce ambiguity, distinguishing between the action of reducing or stopping a drug (discontinuation) and the ensuing signs, symptoms and experiences (withdrawal) seems appropriate.

Relative to intended (therapeutic) drug effects, withdrawal phenomena have been under-researched.[21] Some drug class-specific descriptions exist, but no consensus definition of the physiological and psychological phenomena that may follow dose-reducing or stopping prescribed psychotropics has emerged, excepting those in successive editions of the Diagnostic and Statistical Manual of Mental Disorders.[22] This is likely because (1) complaints similar to withdrawal reactions sometimes precede the prescription, thus challenging the idea that drugs cause such reactions, and (2) mainstream discourse for decades has portrayed psychiatric drugs as conventional medical remedies that target physical abnormalities rather than as psychoactive drugs with "diverse, diffuse, and variable effects on mental life regardless of why they are used".[6,23] Efforts to separate prescribed psychoactive drugs and their users from their illicit but similar counterparts intensified periodically, along with peaks in the War on Drugs and concerns over addiction to illicit substances.[24]

Nonetheless, withdrawal phenomena are well-known in medicine and defined in medical dictionaries as responses or reactions to the cessation or reduction of a substance,[25,26] implying a causal relationship. The American Society of Addiction Medicine defines withdrawal syndrome as "the onset of a predictable constellation of signs and symptoms following the abrupt discontinuation of, or rapid reduction in, the dose of a psychoactive substance".[27] This definition, while useful, may be limited by its emphasis on the first appearance (onset) of symptoms and on abrupt discontinuation. Regarding SSRIs and antipsychotics, Chouinard and colleagues distinguished between onset, duration, outcome, and symptoms of four types of drug withdrawal[28,29]: new withdrawal syndromes and new rebound syndromes (both linked to plausible drug effects on neurotransmitter systems); persistent post-withdrawal disorders; and relapse (recurrence) of the original distress. The inclusion of persistent post-withdrawal disorders – which comprise earlier conceptions of tardive or protracted withdrawal reactions, iatrogenic stress syndromes, and new or unusual presentations of previous symptoms – is an important and belated recognition in mainstream psychiatry of the complexity and reach of withdrawal phenomena.[30–33]

This complexity and reach was long recognized in lay user groups, especially on websites that offered information, guidance, and taper schedules for discontinuing benzodiazepines, antidepressants, and antipsychotics.[34–37] Until very recently, the multiplicity and ingenuity of lay and non-medical perspectives around discontinuing psychotropic drugs to accommodate the user's preferences, and detailed descriptions of the range of withdrawal reactions, contrasted with the dearth of discussions in the psychiatric literature. This may be due partly to the observed emphasis in lay users' descriptions, compared with professional descriptions of drug effects, of "greater context and situational examples of how effects may manifest in various combinations and to varying degrees".[38] In any case, simply reading first-person accounts from disparate users who state that withdrawing from an antidepressant was "definitely the most unpleasant experience of my life" or "the most painful, miserable, and trying experience of my life" should alert anyone that even mild withdrawal symptoms could confound any assessment of "relapse" in a clinical trial.[39,40]

Body, behavior, and environment during withdrawal

Withdrawal effects fit within conceptions of body–drug interactions as tending to adaptation and homeostasis. The neurobiological mode of action of psychotropic drugs remains conceived primarily as the alteration of synaptic neurotransmission by interacting directly, at least initially, with cell membrane-embedded or -associated molecules.[10,41] Drugs alter presynaptic neurons to prompt the release or inhibition of neurotransmitters, they alter postsynaptic neurons by affecting neurotransmitters' binding to receptors, and they alter neurotransmitter reuptake processes and the synthesis of enzymes and receptors and other proteins. Each alteration elicits opposing or compensatory reactions, such as accumulation or depletion of receptors and enzymes, and decreases and increases in levels of associated neurotransmitters. Downstream effects within hours, days, and months extend these alterations to gene expression and transcription and to other neurotransmitter systems, and their ensuing oppositional or compensatory reactions. These responses are believed to allow a certain homeostasis so that

the organism continues to function within a normal range despite the presence of the drug – though it may be, as Galen believed, that a drug ultimately "masters the forces of the body" (p. 11).[42] When the drug is removed, the responses overshoot and the homeostasis is disturbed, leading to the psychobiological discomforts called withdrawal symptoms or syndromes.[28,43–46] In sum, "withdrawal effects are part of the pharmacology of a drug".[47]

However, psychological and social-interpersonal factors, not necessarily reducible to neurobiology, contribute. They may involve players other than the individual who ingests the drug, such as groups of people who exchange ideas. The nocebo effect, or the deterioration of outcomes due to negative expectations, would play a part in withdrawal through verbal suggestions, previous experiences, and vicarious learning or observation.[48] As inferred from animal studies, anticipation of withdrawal symptoms, and other environmental cues, may create dysphoria, which is difficult to distinguish from, and may compound, a drug-induced effect.[49] Incomplete or misleading information given to a patient, such as denial that withdrawal effects could occur, can thwart that individual's ability to interpret and address such symptoms.[39,40] In discontinuation RCTs, attitudes and behavior of study personnel and clinical attendants are sometimes reported to have important effects on the outcomes of (especially institutionalized) patients' discontinuations.[20] First-person accounts also suggest that the economic ability to suspend one's usual duties may determine the success of discontinuation.[50] Knowledge on psychosocial contributions ("set" and "setting") to prescribed drug effects is not fully organized,[51] and rarely do RCTs complement their findings with direct evaluations of treatments by participants in their own words. Still, like Ashton and colleagues, we entertain that psychosocial factors help to produce "the greatest variability, and unpredictability, in discontinuation/withdrawal syndromes ... changes which may be unique to each individual" (p. 297).[52]

Finally, Baldessarini and Tondo point out that no iatrogenic complications associated with discontinuing drugs are measured in a recent epidemiological study of the course of mental disorder (which focuses on the likelihood of being diagnosed with a second mental disorder once a first has been diagnosed).[21] They and others imply that the "natural course of psychiatric illness" has been modified as a result of widespread long-term use and repeated discontinuations of psychotropic drugs,[53] including by most subjects entering a drug trial in developed countries. Their comment acutely raises the question of the ubiquity of "persistent post-withdrawal" effects.

Biases and misconceptions in psychopharmacology RCTs

After decades atop the hierarchy of evidentiary rigor in biomedicine and other fields, the place and reach of the RCT is being re-evaluated,[54] especially its distancing from prior knowledge in psychopharmacology, such as single-case studies, cross-over studies, observational studies, patient narratives, and history of ideas about how treatments work.[55] Outside this "cumulative science," an RCT cannot stand alone.

Before turning to the specific case that withdrawal confounding and its disguise might exist, we emphasize that nearly every strategy potentially favoring the tested drug is built into the design of conventional psychopharmacology RCTs.[56–61] Across different drug classes and treated conditions, a list of these strategies includes:

- Not recruiting random samples, but rather convenience samples, of the defined populations of interest;
- Not recruiting normal volunteers in phase I and II studies to assess psychoactive effects that could account for the outcomes of interest in later studies;
- Excluding individuals likely to be prescribed drugs in practice, especially those with the most disabling problems most likely to justify the prescription of drugs in the real world;
- Excluding individuals with histories of non-response to drug treatment and individuals who respond positively to placebo and/or drug discontinuation during a trial's washout phase;
- Assuming that the somatic disease homogeneity (which justifies randomization to treatment groups) holds where participants meet diagnostic criteria lacking any biological marker detectable in an individual participant;
- Unknown reliance on professional patients;
- Short trial duration compared with real-world duration of drug use;
- Allowing polypharmacy (typically benzodiazepines to induce sleep or reduce agitation

and antiparkinsonians to manage extrapyramidal symptoms) though trials are presented as tests of one drug;

- Employing high doses and unusual schedules of comparator drugs;
- Rarely using active placebos;
- Rarely testing for penetration of the blind;
- Using symptom (outcome) scales sensitive to sedative drug effects or otherwise plagued with validity problems;
- Relying on clinician-rated rather than self-report scales;
- Making and reifying arbitrary distinctions between "responders" and "non-responders" to study drugs;
- Employing definitions of "relapse" not reflecting indicators of a clinically significant relapse;
- Obtaining information about unwanted effects from spontaneous patient comments rather than pointed questioning;
- Classifying and reporting unwanted effects obtusely;
- Not investigating patients' post-treatment ratings or the persistence of drug effects.

In addition to the above recruitment and design strategies, publication or reporting biases further inflate efficacy estimates of tested drugs:[62–66]

- selective publication of entire trials or of undesirable outcomes within trials, and switching trial outcomes from originally stated primary outcomes;
- describing safety assessments with one-quarter the number of words used to describe efficacy assessments;
- duplicate or multiple publication; and
- unattributed authorship (ghostwriting).

Arguably, these strategies turn the average psychopharmacology RCT into a collection of contrived tests and interpretations ill-suited to guide real-life decisions by clinicians and their patients. This impression is reinforced when examining how published reports of RCTs address discontinuing drugs from participants and assessing withdrawal effects.

The relapse-prevention RCT: origins, uses, and biases

Four main types of trials (Table 1) intentionally discontinue tested psychotropic drugs from participants as the key feature of the trial's design.[20] Relapse prevention RCTs are one of these types and our critique does not apply equally to all. First proposed in 1975 by Janssen Pharmaceutica researchers as a way of avoiding undue placebo treatment",[67] they typically start with an open-label period of about 4–12 weeks' duration in which all participants are given the trial drug and assessed for response to this drug (a threshold score on a symptom-rating scale); and a second phase in which non-responders are removed from the trial and responders are randomized to either continue on the same drug (maintenance, of varying length) or be switched to placebo. Maintenance and placebo groups are then compared at the study's end on the same scale; if the maintenance group fares better on average than the placebo group, the drug is said to "prevent relapse."

Ghaemi and Selker argue that relapse-prevention RCTs are tautological in what they purport to measure: participants who respond to drugs in the first phase of the trial are assessed for response to drugs in the second phase (i.e. the predictor and the outcome are the same),[68] which by itself would make results "invariably positive." In another analysis, Hengartner remarks that scientific evidence for long-term antidepressant treatment rests "almost exclusively" on relapse-prevention trials and also observes "unequivocally positive" results of such trials.[69]

The advantages of the relapse-prevention design for non-curative drugs having smallish effects relative to placebo stand out: randomization is applied to a highly selected sample (no placebo responders, no noncompliers, no victims of serious adverse events) after an open-label period where dosages have been optimized. Earlier observers of this design also noted that it overestimated efficacy and underestimated long-term toxicity,[70,71] but its first applications in psychopharmacology were to determine whether a significant withdrawal syndrome occurred upon abrupt discontinuation of benzodiazepines and lithium.[72,73]

We found only a few original trials over the last 30 years that investigated whether withdrawal could present as relapse and was difficult to distinguish from relapse.[74–76] Evaluating the conflicting evidence for antidepressants' long-term effectiveness, Hengartner concluded that "there is substantial withdrawal confounding in

Table 1. Four main types of randomized drug discontinuation designs.

Study type	Rationale	Aim	Design characteristics
Relapse-prevention	Drug efficacy and safety require evaluation.	Use discontinuation to infer conclusions about a drug's efficacy in preventing recurrence of a disorder.	Drug responders during open-label treatment are randomly assigned to continuation or discontinuation groups, whose average outcome scores are compared at study's end.
Successful discontinuation	Long-term drug use produces harm and withdrawal symptoms, and non-drug alternatives are safer.	Help patients discontinue a drug and remain off it.	Long-term drug-treated patients are assigned into groups employing different discontinuation strategies, whose success is determined at study's end.
Practicality of discontinuation	Long-term drug use produces harm, and discontinuation might help.	Determine practicality of drug-free management.	Long-term drug-treated patients are assigned to continuation or discontinuation groups and followed for months while functioning is assessed, then compared.
Architecture of withdrawal	Drug withdrawal syndromes are not well known.	Describe withdrawal symptoms and syndromes.	A drug is discontinued, and ensuing symptoms are noted and related to the discontinuation strategy and patient characteristics.

Adapted from Cohen and Recalt.[20]

discontinuation trials, which renders their findings uninterpretable" (p. 1).[69]

In previous work systematically reviewing 80 discontinuation RCTs published since 2000,[20] 70% with industry funding or participation, we found that, in most RCTs, most of the difference in relapse rates between drug-continued and drug-discontinued groups occurred soon after discontinuation,[77] when withdrawal symptoms are most likely to occur. We identified several flaws in these trials.

Ignoring washout
Washout or placebo lead-in refers to discontinuing trial participants a few days before the trial commences from any drugs taken to ensure a clean slate that avoids carry-over effects. With washout, participants in a relapse-prevention trial might undergo two separate discontinuations with two sets of potential withdrawal reactions. We found that 26.25% of trials reported a washout, with another 20% not reporting one despite being otherwise similar in design and funding source. Almost no information was given on how many participants underwent washout or how they might have been affected. Washout might extend the influence and confound of withdrawal phenomena to any psychiatric drug trial design that employs it.

Not attending to the possibility of a withdrawal confound
As Deaton and Cartwright state in their analysis of the validity of RCTs, "warrant is required that there are no significant post-randomization correlates with the treatment" (p. 9).[54] If withdrawal effects arising post-randomization correlate significantly with outcomes in the group assigned to discontinuation, obviously the trial should acknowledge this possibility or test it. Searching for signs of this attention, we found that over half of discontinuation RCTs showed basic awareness (i.e., brief mention that the drug class might be associated with withdrawal), and 42% included a measure of withdrawal symptoms distinct from extrapyramidal symptoms, but only 13% mentioned that withdrawal might resemble relapse or confound its assessment, and 6% attempted to analyze or account for such a possible confound.

Discontinuing abruptly
We found that nearly 78% of RCTs employed either abrupt (same-day) discontinuations or rapid tapers (under 1 month). Trials of antidepressants and stimulants typically employed abrupt or rapid tapers, trials of antipsychotics occasionally, but those of benzodiazepines never did. Consensus exists that abrupt discontinuations are more likely than slower (over 1 month) methods to induce withdrawal, and to do so

more severely,[28,78] even if more time may not eliminate all withdrawal risk.[79] The prevalence of abrupt or rapid discontinuations is staggering not only considering the risk of biasing study results, but in light of ethical imperatives in medicine and in human subject research to do no harm.

Not attending to the social environment

Surveys and first-person accounts of withdrawal difficulties emphasize support from physicians, therapists, and other people in helping discontinue psychotropic drugs taken for years.[80] Looking only at non relapse-prevention RCTs (i.e., 29 RCTs testing strategies to help individuals discontinue from drugs and stay off them, or examining whether drug-free management of certain dependent or institutionalized individuals might be safe, see Table 1), we observed that 12 did acknowledge environmental, attitudinal, or interpersonal factors – but only two formally assessed them.[81,82] This deficit hints at the lack of relevance to real-world circumstances even of discontinuation trials that do not exploit withdrawal difficulties.

Providing insufficient data to re-analyze outcomes

A vast amount of information is generated in clinical trials that is not described in their published journal reports.[83,84] Of the 80 discontinuation RCTs we reviewed, 85% variously lacked basic information about symptom scores, measures of variability for those scores at important points throughout each trial, information about blinding and concealment of treatment allocation; washout; environmental factors; rationales for specific discontinuation procedures; and whether other psychotropics were concomitantly prescribed during discontinuation.[20,77] This carelessness constitutes another threat to the standing of RCTs as "gold standard" research methods; assessing RCTs solely by the data included in journal articles is like judging an iceberg only by what we see above the waterline.[85]

Suggestions for current trial methodology

Current conventional study designs can account for the risk of withdrawal confounding, first by replacing the abrupt switch to placebo at the start of a randomized discontinuation phase with a slow medication taper of meaningful duration. *Some theoretical justification for any discontinuation*

procedure(s) must be given. In our systematic review, we found that 67% of 77 discontinuation RCTs (including 78% of 23 employing abrupt discontinuation) offered none. The hyperbolic dose-receptor occupancy relationship is one such justification.[86]

Active placebos (substances that mimic some mild, usually physical, effects of the tested drugs) have long been considered a way to raise the bar for tested drugs,[56] but have disappeared from modern efficacy trials in psychiatry, remaining in neurology pain trials. The main argument raised against active placebos is "the risk of unintended therapeutic effect"—an outcome that might be valued in a real-world setting. Not accompanying discontinuation with meaningful support also requires justification.[87] In addition, using measures of withdrawal effects during the discontinuation phase of the trial alongside measures already used to track other effects or states (but rated by different trial assistants) would be a major improvement. Chouinard and colleagues have proposed diagnostic criteria to distinguish between post-discontinuation phenomena,[28,29,88] which could be employed as part of a new willingness to measure potential withdrawal syndromes and factor them into statistical analyses of trial outcomes. The easiest way to do this is to add such measures to regression models (commonly proportional hazards models) as new covariates. Reported measures of relapse risk (RR, OR) or hazard (HR) between drug-continued and discontinued participant groups would thus have controlled for any potential withdrawal confound. Also, assuming that some severe withdrawal symptoms would emerge relatively soon after drug cessation, trialists might reanalyze their main results by excluding participants deemed to have "relapsed" within an interval of time post-discontinuation (based on empirical knowledge about the emergence of withdrawal symptoms for that drug or class). Chen and colleagues employed this reanalysis in their relapse-prevention trial of quetiapine,[89] observing a 32% smaller relapse rate. Some or all of these changes would increase the cost of conducting clinical trials. Finally, addressing the opacity of RCTs requires concrete efforts by trialists and funders to increase transparency and reproducibility, by more study pre-registration, data sharing, and thorough reporting. In 2020, the web can easily be leveraged to provide space for the kinds of detailed description missing in many journal publications (due to space constraints) but essential for RCTs' scientific value.

Limitations

While we assert that outcomes of relapse-prevention RCTs in psychopharmacology are confounded by withdrawal effects, estimating the true magnitude of this confounding is in its infancy. Our 2019 analysis of potential withdrawal confounding was, to our knowledge, the first to do so across published studies of multiple drug classes,[77] but analysis was limited due to missing data in the source publications. Whether trial outcomes would be attenuated, reversed, or unaffected upon more detailed analysis of withdrawal confounding is still unclear.

We judged it beyond the scope of this review to critique procedures in psychopharmacology RCTs that standardize and quantify – typically regarded as the strengths of RCTs – but that, in our view, by relying on surrogate measures from clinicians that distort the experiences of participants, help to paint incomplete pictures of study drugs. We outline many risks of biases, observed biases, and misconceptions in psychopharmacology RCTs and relapse-prevention RCTs specifically, but our suggestions for improvement do not address each of these points, and do not rest on a thorough discussion of using placebo effects in research and in practice contexts. Implementing those suggestions would not rid psychopharmacological RCTs of bias entirely. Instead we hope to encourage the community of researchers in and around psychopharmacology into new ways of thinking. Our broad call for heterodoxy in psychiatric drug trial design and implementation is a call for creativity in addressing this long list of issues, with an accompanying openness to "failure" should a given proposed fix not succeed. We do not see ourselves as capable of providing all the relevant fixes, and this paper is limited in that its critique outweighs its concrete proposals.

We may overstate the impact of non-neurobiological factors (social, environmental) in psychiatric drug action and effects. That the importance of these factors has not been thoroughly investigated, however, underscores our broader point: current psychopharmacology trials work in a narrowly defined way, and potentially useful or enlightening sources of explanation, causality, or therapeutic and adverse effects have gone largely unexamined. We would be happy to be proven wrong on this front if it meant that psychopharmacology and the broader mental health professions and research funders devote more resources to looking at these factors.

Conclusion

Efforts to ensure the validity of experiments are a form of "stacking the deck" against human bias, preference, and interests. Scientists are expected to try in countless ways to ensure that a desirable result stemming from their analyses has run a gauntlet of theoretical, methodological, and analytical checks to ensure that it is true or real. This paper has argued that investigators in psychopharmacology RCTs simply do not stack the deck against the assumption that tested drugs are effective, which is why their efforts to date have failed to confirm that positive results from published drug trials are genuine; that is, they are not reflected in other ways of measuring improvement and recovery, like historical analysis and population-level trends.[15,90] As a result, "an ever-closer adherence to what may appear to be the best evidence could lead to a deterioration in the health of patients" (p. 18).[91] The current period of great difficulty for psychopharmacology research may be called a crisis,[10] and may sit in a larger "replication revolution" across many fields.[92] Addressing it requires vastly increased quality control – stacking the deck against hoped-for outcomes at every stage. While the role of commercial forces in corroding and corrupting this scientific endeavor is undeniable, ridding ourselves of pharmaceutical companies (if such a thing were possible) might not necessarily bring about a valid psychopharmacology that serves those who choose to use drugs and those who choose to help them in this endeavor.

Simpler suggestions directed at medical journals have been ignored,[93] and we are not optimistic that our broad suggestion will be acted upon soon. In any case, orientations to a problem matter greatly to any attempts to address it, and we note efforts to involve patient/consumers at all stages of clinical trial design, including the tapering of psychiatric drugs,[94] although the drug industry may dominate this effort too.[95] We repeat, then, a more modest suggestion we have seen in various forms previously,[96,97] that researchers with contrasting or opposing views (allegiances) about the helpfulness of psychiatric drugs must participate in designing and executing trials together as part of a genuine re-imagining of how we study these basic questions: "Can psychotropic compounds and placebos alleviate psychological distress, and if so, in what contexts and for how long, and what methods are appropriate and ethical to discover this?" We cannot imagine a stronger antidote to

bias than obliged philosophical, cultural, theoretical, and methodological diversity in this common empirical enterprise.

ORCID iDs

David Cohen (iD) https://orcid.org/0000-0002-4970-0804

Alexander Recalt (iD) https://orcid.org/0000-0003-3502-3299

References

1. Geddes JR, Carney SM, Davies C, *et al.* Relapse prevention with antidepressant drug treatment in depressive disorders: a systematic review. *Lancet* 2003; 361: 653–661.

2. Thompson J, Stansfeld JL, Cooper RE, *et al.* Experiences of taking neuroleptic medication and impacts on symptoms, sense of self and agency: a systematic review. *Soc Psychiatry Psychiatr Epidemiol* 2020; 55: 151–164.

3. Read J and Sacia J. Using open questions to understand 650 people's experiences with antipsychotic drugs. *Schiz Bull* 2020; 46: 896–904.

4. Winkelbeiner S, Leucht S, Kane JM, *et al.* Evaluation of differences in individual treatment response in schizophrenia spectrum disorders: a meta-analysis. *JAMA Psychiatry* 2019; 76: 1063–1073.

5. Plöderl M and Hengartner MP. What are the chances for personalized treatment with antidepressants? Detection of patient-by-treatment interaction with a variance ratio meta-analysis. *BMJ Open* 2019; 9: e034816.

6. Jacobs DH and Cohen D. What is really known about psychological alterations produced by psychiatric drugs? *Int J Risk Saf Med* 1999; 12: 37–47.

7. Wipond R. Are psychiatric medications safe? The FDA's answer may surprise you. *Inner compass initiative*, https://www.theinnercompass.org/blog/are-psychiatric-medications-safe-fdas-answer-may-surprise-you (2020, accessed 5 June 2020).

8. Kirk SA, Gomory T and Cohen D. *Mad science: psychiatric coercion, diagnosis, and drugs*. 1st ed. New York: Routledge, 2013.

9. Zachar P, Stoyanov DS, Aragona M, *et al.* (eds). *Alternative perspectives on psychiatric validation*. Oxford: Oxford University Press, 2015.

10. Drukarch B, Jacobs GE and Wilhelmus MMM. Solving the crisis in psychopharmacological research: cellular-membrane(s) psychopharmacology to the rescue? *Biomed Pharmacother* 2020; 130: 110545.

11. Healy D and Mangin D. Clinical judgments, not algorithms, are key to patient safety. *BMJ* 2019; 367: l5777.

12. Lewis TR, Reichman JH and So AD. The case for public funding and public oversight of clinical trials. *Economists' Voice* 2007; 4: 1–4.

13. Moore TJ and Mattison DR. Adult utilization of psychiatric drugs and differences by sex, age, and race. *JAMA Intern Med* 2017; 177: 274–275.

14. Lopez-Leon S, Lopez-Gomez MI, Warner B, *et al.* Psychotropic medication in children and adolescents in the United States in the year 2004 vs 2014. *Daru* 2018; 26: 5–10.

15. Insel T. Translating scientific opportunity into public health impact. *Arch Gen Psychiatry* 2009; 66: 128–133.

16. Kane JM and Lieberman JA. (eds). *Adverse effects of psychotropic drugs*. New York: Guilford Press, 1992.

17. Breggin PR and Cohen D. *Your drug may be your problem: how and why to stop taking psychiatric medication*. Philadelphia, PA: Perseus Books, 1999.

18. Nielsen M, Hansen EH and Gøtzsche PC. Dependence and withdrawal reactions to benzodiazepines and selective serotonin reuptake inhibitors. How did the health authorities react? *Int J Risk Safety Med* 2013; 25: 155–168.

19. Fava GA, Gatti A, Belaise C, *et al.* Withdrawal symptoms after selective serotonin reuptake inhibitor discontinuation: a systematic review. *Psychother Psychosom* 2015; 84: 72–81.

20. Cohen D and Recalt AR. Discontinuing psychotropic drugs from participants in randomized controlled trials: a systematic review. *Psychother Psychosom* 2019; 88: 96–104.

21. Baldessarini RJ and Tondo L. Effects of treatment discontinuation in clinical psychopharmacology. *Psychother Psychosom* 2019; 88: 65–70.

22. American Psychiatric Association. *Diagnostic and statistical manual of mental disorders*. 5th ed.

Washington, DC: American Psychiatric Press, 2013.

23. Moncrieff J and Cohen D. How do psychiatric drugs work? *BMJ* 2009; 338: b1963.

24. Herzberg D. Entitled to addiction? Pharmaceuticals, race, and America's first drug war. *Bull Hist Med* 2017; 91: 586–623.

25. Mosby. *Mosby's medical dictionary*. St. Louis, MI: Elsevier Health Sciences, 2016.

26. Venes D. *Taber's cyclopedic medical dictionary*. Philadelphia, PA: FA Davis, 2017.

27. Ries RK, Fiellin DA, Miller SC, *et al. The ASAM principles of addiction medicine*. Chevy Chase, MD: American Academy of Addiction Medicine; Philadelphia, PA: Wolters Kluwer, 2014.

28. Chouinard G and Chouinard V-A. New classification of serotonin reuptake inhibitor withdrawal. *Psychother Psychosom* 2015; 84: 63–71.

29. Chouinard G, Samaha A-N, Chouinard V-A, *et al.* Antipsychotic induced dopamine supersensitivity psychosis: pharmacology, criteria, and therapy. *Psychother Psychosom* 2017; 86: 189–219.

30. Ashton H. Protracted withdrawal syndromes from benzodiazepines. *J Subst Abuse Treat* 1991; 8: 19–28.

31. Lewander T. Neuroleptics and the neuroleptic-induced deficit syndrome. *Acta Psychiatrica Scand* 1994; 380: 8–13.

32. Healy D and Tranter R. Pharmacological stress diathesis syndromes. *J Psychopharmacol* 1999; 13: 287–290.

33. Belaise C, Gatti A, Chouinard V-A, *et al.* Patient online report of selective serotonin reuptake inhibitor-induced persistent post-withdrawal anxiety and mood disorders. *Psychother Psychosom* 2012; 81: 386–388.

34. Everything Matters Beyond Meds. Psychiatric drug withdrawal 101, https://beyondmeds.com/withdrawal-101/ (accessed 5 June 2020).

35. Surviving Antidepressants. https://www.survivingantidepressants.org (accessed 5 June 2020).

36. The Withdrawal Project. Inner compass initative, https://withdrawal.theinnercompass.org (accessed 5 June 2020).

37. Mad in America. MIA's drug withdrawal resources, https://www.madinamerica.com/drugs-withdrawal-home/ (accessed 5 June 2020).

38. Hughes S and Cohen D. Can online consumers contribute to drug knowledge? A mixed methods comparison of consumer-generated and professionally controlled psychotropic medication information on the internet. *J Med Internet Res* 2011; 13: e53.

39. Simons P. Peer-support groups were right, guidelines were wrong: Dr. Mark Horowitz on tapering off antidepressants. *Mad in America*, https://www.madinamerica.com/2019/03/peer-support-groups-right-official-guidelines-wrong-dr-mark-horowitz-tapering-off-antidepressants/ (2019, accessed 5 June 2020).

40. Blumke D. Ambushed by antidepressant withdrawal: the escape story. *Mad in America*, https://www.madinamerica.com/2019/09/ambushed-antidepressant-withdrawal/ (2019, accessed 5 June 2020).

41. Stahl SM. *Stahl's essential psychopharmacology: neuroscientific basis and practical applications*. 3rd ed. Cambridge: Cambridge University Press, 2008.

42. Escohotado A. *The general history of drugs*. Vol. 1. Trans. Robinette GW. Valparaiso, Chile: Graffiti Militante Press, 2010.

43. Koob GF and Le Moal M. *Neurobiology of addiction*. San Diego, CA: Academic Press, 2006.

44. Fava GA and Offidani E. The mechanisms of tolerance in antidepressant action. *Prog Neuro-Psychopharmacol Biol Psychiatry* 2011; 35: 1593–1602.

45. Andrews PW, Kornstein SG, Halberstadt LJ, *et al.* Blue again: perturbational effects of antidepressants suggest monoaminergic homeostasis in major depression. *Front Psychol* 2011; 2: 159.

46. Lerner A and Klein M. Dependence, withdrawal, and rebound of CNS drugs: an update and regulatory considerations for new drugs development. *Brain Commun* 2019; 1: fcz025.

47. Reidenberg MM. Drug discontinuation effects are part of the pharmacology of a drug. *J Pharmacol Exp Ther* 2011; 339: 324–328.

48. Colloca L. Nocebo effects: the dilemma of disclosing adverse events. In: Strech D and Mertz M (eds) *Ethics and governance of biomedical research: theory and practice*. Switzerland: Springer, 2016.

49. Siegel S. Drug anticipation and the treatment of dependence. *NIDA Res Monogr* 1988; 84: 1–24.

50. Aviv R. The challenge of going off psychiatric drugs. *The Atlantic*, https://www.newyorker.com/magazine/2019/04/08/the-challenge-of-going-off-psychiatric-drugs (2019, accessed 3 June 2020).

51. Hartogsohn I. Constructing drug effects: a history of set and setting. *Drug Sci Policy Law* 2017; 3: 1–17.

52. Ashton H, Young AH and Ferrier N. Psychopharmacology revisited. *J Psychopharmacol* 1999; 13: 296–298.

53. Fava GA and Rafanelli C. Iatrogenic factors in psychopathology. *Psychother Psychosom* 2019; 89: 129–140.

54. Deaton A and Cartwright N. Understanding and misunderstanding randomized controlled trials. *Soc Sci Med* 2018; 210: 2–21.

55. Rocca E and Janum RL. Causal evidence and dispositions in medicine and public health. *Int J Environ Res Public Health* 2020; 17: 1813.

56. Fisher S and Greenberg RP. What are we to conclude about psychoactive drugs? Scanning the major findings. In: Fisher S and Greenberg RG (eds) *From placebo to panacea: putting psychiatric drugs to the test*. New York: John Wiley & Sons, 1997, pp.359–384.

57. Cohen D. Drug treatment of schizophrenia: a critical appraisal and implications for social work education. *J Soc Work Educ* 2002; 38: 217–239.

58. Cohen D. Clinical psychopharmacology trials: gold standard or fool's gold? In:Kirk SA (ed.) *Mental disorders in the social environment*. New York: Columbia University Press, 2005, pp.347–67.

59. Cohen D and Jacobs D. Randomized controlled trials of antidepressants: clinically and scientifically irrelevant. *Deb Neurosci* 2007; 1: 44–54.

60. Ioannidis JP. Effectiveness of antidepressants: an evidence myth constructed from a thousand randomized trials? *Philos Ethics Humanit Med* 2008; 3: 14.

61. Moncrieff JM, Crellin JM, Long MA, *et al.* Definitions of relapse in trials comparing antipsychotic maintenance with discontinuation or reduction for schizophrenia spectrum disorders: a systematic review. *Schiz Res* 2019. Epub ahead of print 8 October 2019. DOI: 10.1016/j.schres.2019.08.035.

62. Turner EH. Publication bias, with a focus on psychiatry: causes and solutions. *CNS Drugs* 2013; 27: 457–468.

63. Ioannidis JP. The mass production of redundant, misleading, and conflicted systematic reviews and meta-analyses. *Millbank Q* 2016; 94: 485–514.

64. Healy D and Cattell D. Interface between authorship, industry, and science in the domain of therapeutics. *Br J Psychiatry* 2003; 183: 22–27.

65. Hughes S, Cohen D and Johnson R. Adverse event assessment methods in published trials of psychotropic drugs: poor reporting and neglect of emerging safety concerns. *Int J Risk Saf Med* 2016; 28: 101–114.

66. Hughes S, Cohen D and Jaggi R. Differences in reporting serious adverse events in industry sponsored clinical trial registries and journal articles on antidepressant and antipsychotic drugs: a cross-sectional study. *BMJ Open* 2014; 4: e005535.

67. Amery W and Dony J. A clinical trial design avoiding undue placebo treatment. *J Clin Pharmacol* 1975; 15: 674–679.

68. Ghaemi SN and Selker HP. Maintenance efficacy designs in psychiatry: randomized discontinuation trials–enriched but not better. *J Clin Transl Sci* 2017; 1: 198–204.

69. Hengartner MP. How effective are antidepressants for depression over the long term? A critical review of relapse prevention trials and the issue of withdrawal confounding. *Ther Adv Psychopharmacol* 2020; 10: 2045125320921694.

70. Friedman LM, Furberg CD, DeMets DL, *et al. Fundamentals of clinical trials*. 2nd ed. Littleton, MA: PSG Publishing Company, 1985.

71. Kopec JA, Abrahamowicz M and Esdaile JM. Randomized discontinuation trials: utility and efficiency. *J Clin Epidemiol* 1993; 46: 959–971.

72. Rickels K, Case WG and Downing RW. Long-term diazepam therapy and clinical outcome. *JAMA* 1983: 250: 767–771.

73. Mander AJ and Loudon JB. Rapid recurrence of mania following abrupt discontinuation of lithium. *Lancet* 1988; 2: 15–17.

74. Greenhouse JB, Stangl D, Kupfer DJ, *et al.* Methodologic issues in maintenance therapy clinical trials. *Arch Gen Psychiatry* 1991; 48: 313–318.

75. Goodwin GM, Emsley R, Rembry S, *et al.* Agomelatine prevents relapse in patients with major depressive disorder without evidence of a discontinuation syndrome: a 24-week randomized, double-blind, placebo controlled trial. *J Clin Psychiatry* 2009; 70: 1128–1137.

76. Sunder KR, Wisner KL, Hanusa BH, *et al.* Postpartum depression recurrence versus discontinuation syndrome: observations from

a randomized controlled trial. *J Clin Psychiatry* 2004: 65: 1266–1268.

77. Récalt AM and Cohen D. Withdrawal confounding in randomized controlled trials of antipsychotic, antidepressant, and stimulant drugs, 2000–2017. *Psychother Psychosom* 2019; 88: 105–113.

78. Fava GA, Bernardi M, Tomba E, *et al.* Effects of gradual discontinuation of selective serotonin reuptake inhibitors in panic disorder with agoraphobia. *Int J Neuropsychopharmacol* 2007; 10: 835–838.

79. Groot PC and van Os J. Antidepressant tapering strips to help people come off medication more safely. *Psychosis* 2018; 10: 142–145.

80. Ostrow L, Jessell L, Hurd M, *et al.* Discontinuing psychiatric medications: a survey of long-term users. *Psychiatr Serv* 2017; 68: 1232–1238.

81. Ahmed Z, Fraser W, Kerr MP, *et al.* Reducing antipsychotic medication in people with a learning disability. *Br J Psychiatry* 2000; 176: 42–46.

82. de Kuijper GM, Evenhuis H, Minderaa RB, *et al.* Effects of controlled discontinuation of long-term used antipsychotics for behavioural symptoms in individuals with intellectual disability. *J Intellect Disab Res* 2014; 58: 71–83.

83. Mathieu S, Boutron I, Moher D, *et al.* Comparison of registered and published primary outcomes in randomized controlled trials. *JAMA* 2009; 302: 977–984.

84. Weiseler B, Kerekes MF, Vervoelgyi V, *et al.* Impact of document type on reporting quality of clinical drug trials: a comparison of registry reports, clinical study reports, and journal publications. *BMJ* 2012; 344: d8141.

85. Doshi P, Dickersin K, Healy D, *et al.* Restoring invisible and abandoned trials: a call for people to publish the findings. *BMJ* 2013; 346: f2865.

86. Horowitz MA and Taylor D. Tapering of SSRI treatment to mitigate withdrawal symptoms. *Lancet Psychiatry* 2019; 6: 538–546.

87. Jensen JK, Bielefeldt AO and Hróbjartsson A. Active placebo control groups of pharmacological interventions were rarely used but merited serious consideration: a methodological overview. *J Clin Epidemiol* 2017; 87: 35–46.

88. Cosci F and Chouinard G. Acute and persistent withdrawal syndromes following discontinuation of psychotropic medications. *Psychother Psychosom* 2020; 89: 283–306.

89. Chen EYH, Hui CLM, Lam MML, *et al.* Maintenance treatment with quetiapine versus discontinuation after one year of treatment in patients with remitted first episode psychosis: randomised controlled trial. *BMJ* 2010; 341: c4024.

90. Hegarty J, Baldessarini RJ, Tohen M, *et al.* One hundred years of schizophrenia: a meta-analysis of the outcome literature. *Am J Psychiatry* 1994; 151: 1409–1416.

91. Healy D. Trussed in evidence: ambiguities at the interface of clinical evidence and clinical practice. *Transcult Psychiatry* 2009; 46: 16–37.

92. Gelman A. The competing narratives of scientific revolution. *Statistical modeling, causal inference, and social science,* https://statmodeling. stat.columbia.edu/2018/08/20/competing-narratives-scientific-revolution (accessed 23 August 2020).

93. Smith R. Medical journals are an extension of the marketing arm of pharmaceutical companies. *PLoS Med* 2005; 2: e138.

94. Groot PC and van Os J. How user knowledge of psychotropic drug withdrawal resulted in the development of person-specific tapering medication. *Ther Adv Psychopharmacol* 2020; 10: 2045125320932452.

95. Hansen M, Nørgaard LS and Hallgren CE. How and why to involve patients in drug development: perspectives from the pharmaceutical industry, regulatory authorities, and patient organizations. *Ther Innov Regul Sci* 2020; 54: 577–585.

96. Klein DF. Preventing hung juries about therapy studies. *J Consult Clin Psychol* 1996; 64: 81–87.

97. Quitkin FM, Rabkin JG, Gerald J, *et al.* Validity of clinical trials of antidepressants. *Am J Psychiatry* 2000; 157: 327–337.