

Clinical research challenges posed by difficult-to-treat depression

A. John Rush^{1,2,3} , Harold A. Sackeim⁴, Charles R. Conway⁵, Mark T. Bunker⁶, Steven D. Hollon⁷, Koen Demyttenaere^{8,9}, Allan H. Young¹⁰, Scott T. Aaronson¹¹, Maxine Dibué^{12,13}, Michael E. Thase¹⁴ and R. Hamish McAllister-Williams^{15,16}

Review Article

Cite this article: Rush AJ *et al* (2022). Clinical research challenges posed by difficult-to-treat depression. *Psychological Medicine* **52**, 419–432. <https://doi.org/10.1017/S0033291721004943>

Received: 30 June 2021

Revised: 2 November 2021

Accepted: 9 November 2021

First published online: 7 January 2022

Keywords:

Classification; difficult-to-treat depression; outcome measures; study design; taxonomy; treatment-resistant depression

Author for correspondence:

A. John Rush,

E-mail: curbstoneconsultant@gmail.com

¹Duke-NUS Medical School, Singapore; ²Department of Psychiatry and Behavioral Sciences, Duke University, Durham, NC, USA; ³Department of Psychiatry, Texas Tech University, Permian Basin, TX, USA; ⁴Departments of Psychiatry and Radiology, Columbia University, New York, NY, USA; ⁵Department of Psychiatry, Washington University in St. Louis, St. Louis, MO, USA; ⁶LivaNova USA PLC, Houston, TX, USA; ⁷Departments of Psychology and Psychiatry, Vanderbilt University, Nashville, TN, USA; ⁸University Psychiatric Center, KU Leuven, Leuven, Belgium; ⁹Faculty of Medicine, KU Leuven, Leuven, Belgium; ¹⁰Department of Psychological Medicine, Institute of Psychiatry, Psychology and Neuroscience, King's College, London, UK; ¹¹Department of Clinical Research, Sheppard Pratt Health System, Baltimore, MD, USA; ¹²Department of Neurosurgery, Heinrich Heine University Düsseldorf, Düsseldorf, Germany; ¹³Medical Affairs Europe, LivaNova Deutschland GmbH, Munich, Germany; ¹⁴Department of Psychiatry, University of Pennsylvania, Philadelphia, PA, USA; ¹⁵Northern Centre for Mood Disorders, Newcastle University, Newcastle upon Tyne, UK and ¹⁶Cumbria, Northumberland, Tyne and Wear NHS Foundation Trust, Newcastle upon Tyne, UK

Abstract

Approximately one-third of individuals in a major depressive episode will not achieve sustained remission despite multiple, well-delivered treatments. These patients experience prolonged suffering and disproportionately utilize mental and general health care resources. The recently proposed clinical heuristic of 'difficult-to-treat depression' (DTD) aims to broaden our understanding and focus attention on the identification, clinical management, treatment selection, and outcomes of such individuals. Clinical trial methodologies developed to detect short-term therapeutic effects in treatment-responsive populations may not be appropriate in DTD. This report reviews three essential challenges for clinical intervention research in DTD: (1) how to define and subtype this heterogeneous group of patients; (2) how, when, and by what methods to select, acquire, compile, and interpret clinically meaningful outcome metrics; and (3) how to choose among alternative clinical trial design options to promote causal inference and generalizability. The boundaries of DTD are uncertain, and an evidence-based taxonomy and reliable assessment tools are preconditions for clinical research and subtyping. Traditional outcome metrics in treatment-responsive depression may not apply to DTD, as they largely reflect the only short-term symptomatic change and do not incorporate durability of benefit, side effect burden, or sustained impact on quality of life or daily function. The trial methodology will also require modification as trials will likely be of longer duration to examine the sustained impact, raising complex issues regarding control group selection, blinding and its integrity, and concomitant treatments.

Introduction

Over one-third of persons with major depressive disorder (MDD) will not achieve sustained symptom remission after several treatment trials (Agency for Healthcare Research and Policy (AHQR), 2011). The Sequenced Treatment Alternatives to Relieve Depression (STAR*D) trial found that only two-thirds of patients reached remission after four treatment steps. Furthermore, the relapse rate over the year subsequent to remission ranged from 35% to 70%, increasing with the number of acute treatment trials needed to achieve remission (Rush *et al.*, 2006b).

These and other findings formed the basis for the heuristic, treatment-resistant depression (TRD), which is typically defined by the number of previously failed acute phase treatment trials, based on lack of short-term improvement in overall depressive symptom severity (Fava, 2003b; Sackeim, 2001; Thase & Rush, 1995). Various reports have used thresholds ranging from 1 to 4 or more failed acute phase treatment trials to define various levels of treatment resistance (Agency for Healthcare Research and Policy (AHQR), 2011; Conway, George, & Sackeim, 2017; Lisanby *et al.*, 2009). While the FDA recognizes the utility of TRD in its approval and labeling of interventions, the heuristic poses a myriad of clinical and research challenges (McAllister-Williams *et al.*, 2020; Rush, Aaronson, & Demyttenaere, 2019).

Favoring the concept of TRD is the fact that the degree of resistance seems to be easily assessed (Berlim & Turecki, 2007b; Sackeim *et al.*, 2019). Generally, the likelihood of response

© The Author(s), 2022. Published by Cambridge University Press. This is an Open Access article, distributed under the terms of the Creative Commons Attribution licence (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted re-use, distribution and reproduction, provided the original article is properly cited.

or remission with subsequent treatments decreases with increasing numbers of previously failed acute phase treatment trials (Heijnen, Birkenhager, Wierdsma, & van den Broek, 2010; Lisanby *et al.*, 2009; Prudic *et al.*, 1996; Rush *et al.*, 2006b), while the likelihood of relapse increases if short-term remission is obtained (Prudic *et al.*, 2013; Rasmussen *et al.*, 2009; Rush *et al.*, 2006b; Sackeim *et al.*, 1990).

On the other hand, closer examination of the concept of TRD presents challenges. For example, what defines a failed trial: lack of response or lack of remission? What constitutes an adequate trial? What about patients who markedly improve but do not stay better? Or those who cannot tolerate a medication (is that a failed trial)? Must all the failed trials occur in the current episode or do failed treatments in prior episodes also count, especially since a treatment that failed in the past is expected to fail again if tried in a new episode? All of these questions can be operationalized and attempts have been made to do this on the basis of a Delphi consensus approach (Sforzini, 2021). Nevertheless, there is inevitably great diversity in the prior 'failed' treatments in individuals with TRD, such that there is no expectation of biological or etiological homogeneity in any group defined solely by TRD. Indeed, definitions of TRD generally don't include non-pharmacological treatments and rarely, if ever, include psychotherapy or psychosocial interventions. Consequently, the treatment implications of TRD are largely nonspecific (i.e. more prior failed trials reduce hopefulness about future therapeutics). Finally, TRD has no practical, actionable clinical implications other than to suggest attempting another, primarily pharmacological, treatment trial with a different intervention or combination.

In recognition of these limitations, a new clinical heuristic, difficult-to-treat depressions (DTDs), has been proposed to stimulate the timely identification and personalized management of patients for whom our current treatments – even if well-delivered and tolerated – are unlikely to either initiate or sustain symptomatic remission (McAllister-Williams *et al.*, 2020; Rush *et al.*, 2019). Patients with DTD often present with either chronic depressive symptoms that are insufficiently relieved by treatment changes, or with symptoms that seemingly improve, at least temporarily, but the sustained benefit is not achieved. In either case, persons with DTD have substantially impaired daily function and poor quality of life (QoL) (Jaffe, Rive, & Denee, 2019; McAllister-Williams *et al.*, 2020; Rush *et al.*, 2019). They are also high utilizers of mental and general health services in both outpatient and inpatient settings (Kubitz, Mehra, Potluri, Garg, & Cossrow, 2013; Olchanski *et al.*, 2013), resulting in high health costs that often persist for years (Amos *et al.*, 2018; Benson, Szukis, Sheehan, Alphas, & Yuce, 2020; Greenberg, Corey-Lisle, Birnbaum, Marynchenko, & Claxton, 2004; Olfson, Amos, Benson, McRae, & Marcus, 2018; Sussman, O'Sullivan, Shah, Olfson, & Menzin, 2019; Wang *et al.*, 2005). DTD is also associated with substantial morbidity and mortality (Amital *et al.*, 2013; Huang *et al.*, 2020). Clinical trial findings and care system database analyses (Eaton *et al.*, 2008; Huang *et al.*, 2020; Jaffe *et al.*, 2019; Sussman *et al.*, 2019) suggest that approximately 15–25% of depressed patients present with DTD, though its true prevalence and diagnostic boundaries remain uncertain.

A recent international consensus report detailed the clinical features and treatment implications of DTD, and suggested key principles for management (McAllister-Williams *et al.*, 2020). When DTD is suspected, psychiatric, medical, and neuropsychological re-evaluations are recommended to identify potentially treatable causes of the depressive episode. If the depression is

not improved following these efforts, DTD is 'confirmed' and the treatment goals might need to shift from the pursuit of symptomatic remission to optimizing symptom control, maximizing psychosocial function and QoL, and reducing the risk of deterioration and relapse (McAllister-Williams *et al.*, 2020; Rush *et al.*, 2019), while keeping an eye out for newly developing treatments that may be useful for the patient.

The concept of DTD acknowledges that our therapeutic armamentarium does not presently achieve sustained symptom remission in a significant proportion of depressed patients. Our limited armamentarium may be due, in part, to the heavy reliance on short-term (6–12 week) trials in 'treatment-responsive' populations when developing therapeutics for major depressive episodes (MDEs). These trials typically compare an antidepressant against a placebo, sham, or active comparator, and a statistically and clinically meaningful acute antidepressant effect is potentially identified. Subsequently, continuation or maintenance phase trials address prevention of relapse or recurrence (Frank *et al.*, 1991; Rush *et al.*, 2006a). However, these brief, symptom-focused approaches may be of limited relevance in DTD, as the pursuit of symptom remission through the administration of sequential monotherapies and treatment combinations delivered in a 'try and try again' approach typically results in diminishing returns and may seem futile to the clinician and patient (DeRubeis *et al.*, 2020; Dunner *et al.*, 2006; Hollon *et al.*, 2014; Rush *et al.*, 2006b).

At some point, clinical strategies useful in responsive populations are no longer optimal in DTD. This notion is not unique to DTD. In epilepsy, failure to benefit from two well-conducted trials of anticonvulsant medications is often viewed as the threshold for the diagnosis of medication-resistant epilepsy, triggering additional evaluations and potential surgical intervention (Jette, Reid, & Wiebe, 2014; Kwan & Brodie, 2010). These patients have only an estimated 3–5% chance of achieving at least 1 year of seizure remission with additional antiepileptic medication treatment, and seizure recurrence is common in those who achieve remission (Brodie, Barry, Bamagous, Norrie, & Kwan, 2012; Callaghan, Anand, Hesdorffer, Hauser, & French, 2007). Similarly, in STAR*D, after two failed treatment trials, the probability of achieving remission in MDD dropped 50% in the next two pharmacotherapy steps. Critically, similar to epilepsy, more failed acute-phase antidepressant medication trials result in both lower acute response/remission rates and higher rates of relapse during follow-up.

Thus, at some point, clinicians must decide whether to pursue DTD remission with another treatment trial or to change the aim of treatment to optimized symptom control, function, QoL, relapse mitigation, and treatment burden. This decision entails shared decision-making while considering each patient's aspirations, disease and treatment burdens (e.g. medical fragility), environmental circumstances, and anticipated risks and benefits of untried and sometimes minimally evaluated treatments, as well as other factors. This problem of when to change targets from sustained symptom remission to optimal patient and disease management is implicit in managing every depressed patient who is struggling despite multiple treatment attempts. The designation of DTD promotes a more deliberate, transparent and shared decision-making process. Further, it implicitly promotes the use of evidence-based interventions before embarking on less rigorously evidenced treatments.

There is a great need for clinical and service researchers to develop and evaluate patient-level clinical interventions for

DTD, such as novel medications or combinations, psychotherapies, and neuro-stimulatory methods, as well as administrative/programmatic innovations such as intensive outpatient programs, peer-supported self-help programs aimed at promoting wellbeing, or substance use reduction programs.

To facilitate DTD intervention research, this report reviews three fundamental challenges: participant selection, outcome assessment, and study design. Participant selection is especially challenging because DTD is likely heterogeneous in etiology, course, pathobiology, and intervention responsiveness. The development of a DTD taxonomy could facilitate the identification of more homogeneous subgroups, thereby enhancing trial efficiency and, potentially, intervention targeting and staging (Sackeim, 2021). Outcome assessment concerns the selection of primary and secondary outcomes from multiple possibilities and how to efficiently collect, compile, and interpret these outcomes. Study design challenges include how to select among designs that optimize generalizability, while also preserving opportunities for making causal inference, selection of control conditions, study duration, and other issues.

Challenges in identifying DTD patients for intervention trials

What are the preferred evaluations when DTD is suspected?

The consensus report recommended that a broad set of evaluations be considered when DTD is suspected (McAllister-Williams et al., 2020) to identify overlooked modifiable causes of the depression. For example, depression may be a manifestation of undiagnosed endocrine disorders [e.g. hypothyroidism (Duntas & Maillis, 2013; Hage & Azar, 2012) or Cushing's syndrome (Arnaldi et al., 2003; Pivonello et al., 2016; Sonino, Fava, Raffi, Boscaro, & Fallo, 1998)]. However, beyond clinical consensus, there is little empirical guidance on the relative costs and yield of these potential diagnostic tests and procedures. Relevant issues include whether assessment algorithms are specified by symptomatic presentation, treatment history, or sociodemographic characteristics. Similarly, evidence is currently incomplete regarding the potential of pharmacogenetic/genomic testing to identify or subgroup DTD (Vittengl, Clark, Thase, & Jarrett, 2019; Zeier et al., 2018). How these issues are resolved will impact the inclusion and exclusion criteria used by intervention researchers to define DTD.

Define the boundaries and develop a taxonomy for DTD

The current clinically based characterization of DTD (Gaynes et al., 2018, 2020) lacks sufficient specificity to define intervention study subpopulations. The challenges to defining DTD subpopulations are complex: persons with DTD are heterogeneous in treatment history and responsiveness, sensitivity to treatments, prognosis, and pathobiology (McAllister-Williams et al., 2020; Rush et al., 2019). Clinical research with DTD would benefit from an evidence-based taxonomy that can identify more homogeneous subgroups. Such a taxonomy would make intervention research more cost-efficient and potentially improve our ability to match specific interventions with DTD subgroups. These empirically defined subgroups would also assist mechanistic researchers in elucidating the various pathobiological pathways that likely underlie different types of DTD.

To illustrate the need for a DTD taxonomy, consider the STAR*D findings that acute response and remission rates decreased with increasing numbers of failed treatment trials (e.g. 48.6%, 28.5%, 16.8%, 16.3% for the first four acute phase treatment attempts, respectively). In addition, when remission was achieved, relapse/recurrence rates during follow-up increased progressively with more previously failed acute phase treatments (Rush et al., 2006b; Sackeim, 2016). It is unknown whether these two findings reflect unitary or distinct neurobiological effects linked to antidepressant treatment resistance. Indeed, within DTD, some individuals show minimal treatment responsiveness despite repeated interventions, while others experience substantial short-term benefit, but do not stay well. Are these etiologically distinct groups that may require different treatment strategies? A similar concern can be raised when considering individuals with a chronic course with few or no intervening periods of wellness compared to individuals with recurrent, relapsing depression. A course of illness seems to differentially relate to acute and longer-term treatment outcomes in depression (Rush et al., 2012).

Our current working taxonomy for DTD is based on the designation, TRD, which itself is highly variable in operationalization (Berlim & Turecki, 2007a; Gaynes et al., 2018, 2020). TRD is typically ascribed after two unsuccessful, but well-delivered, acute phase treatments, a threshold that is supported by the marked decrease in sustained remission rates in STAR*D after the first two treatment trials (Conway et al., 2017; Rush et al., 2006b).

This empirical model of TRD, however, is not an adequate taxonomy for DTD. Any classification/treatment algorithm of DTD must consider multiple DTD variations. For example, DTD patients who are hypersensitive to medication side effects and cannot tolerate two medication trials may be difficult-to-treat, but are not 'treatment resistant'. DTD patients who receive interventions and show marked transitory benefit which is not sustained are not considered 'treatment resistant' (Gaynes et al., 2018; Sackeim et al., 2019), but clearly are difficult to treat. In addition, the nature of the two failed treatments is not specified (e.g. two selective serotonin reuptake inhibitors (SSRIs) v. one SSRI followed by transcranial magnetic stimulation) (Fava, 2003b; Fekadu et al., 2009), and it is unlikely that all treatments are equivalent in their prognostic implications for insufficient or short-lived benefit.

This dilemma could be addressed, in part, by staging the 'level of resistance' (analogous to cancer staging) (Conway et al., 2017; Thase & Rush, 1995), based either on a minimum number of *specific treatment types* (e.g. monoamine reuptake inhibitors; brain stimulation treatments, depression-targeted psychotherapies) or *specific treatment sequences* [e.g. cognitive behavioral therapy to SSRI to atypical antipsychotic augmentation to monoamine oxidase inhibitor to electroconvulsive therapy (ECT)]. Even with this approach, the introduction of a new treatment will change sample definitions. Further, such staging approaches focus only on treatment history rather than the broader clinical context and course.

Characterizing DTD: Clinical features, perpetuating factors, and temporal evolution

A more fruitful approach may be a multidimensional characterization of DTD based on its associated clinical presentations, clinical course, biomedical, prognostic, neuropsychological, treatment response, and other features. The degree and duration of

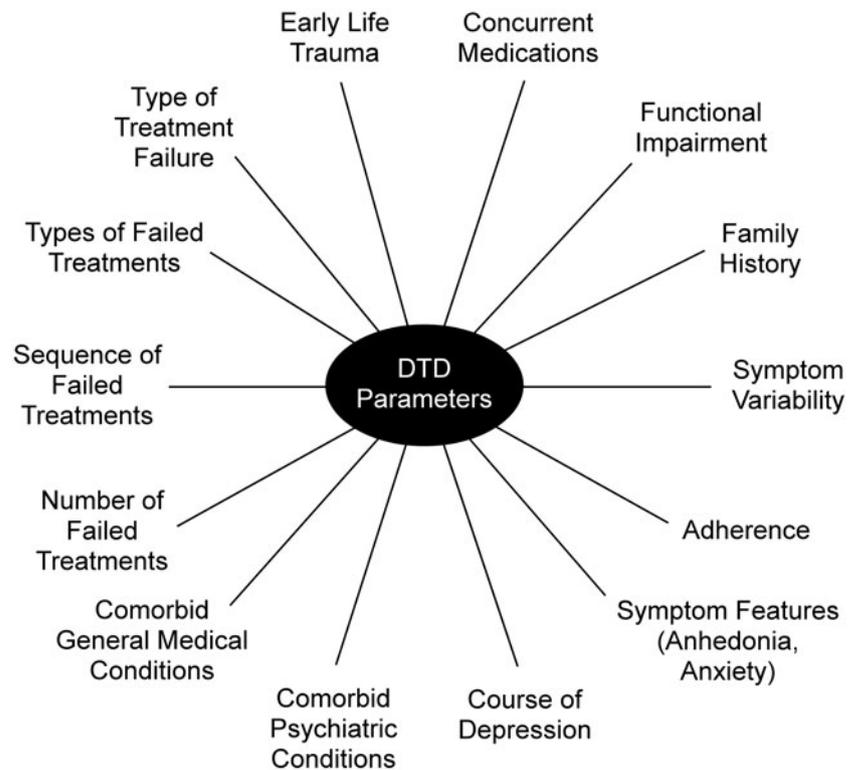


Fig. 1. Potential parameters to define DTD or to characterize subgroups.

functional impairment, history of symptomatic improvement and relapse, and presenting symptoms may be as consequential in identifying and characterizing DTD as the number and type of unsuccessful treatment attempts. These parameters would form the basis for an evidence-based DTD taxonomy that would not necessarily change as new treatments or management tools arrive. These features would define the overall DTD population, and inform the identification of subgroups or spectrums (with or without the addition of biomarkers or pharmaco-dissection). This effort could begin by evaluating whether specific features distinguish DTD from non-DTD patient groups (e.g. concurrent general medical problems; types and severity of anxiety symptoms; types, severity and chronicity of environmental stressors; substance use/abuse disorder; history of childhood trauma/abuse; etc.) (McAllister-Williams et al., 2020; Rush et al., 2019) (Fig. 1).

DTD raises additional challenges to taxonomy development. It is likely that psychosocial determinants of health are perpetuating factors, and these must be addressed to achieve better outcomes, including chronic occupational, marital, economic, or health stressors; co-morbid substance misuse; sedentary lifestyle; obesity; etc. In addition, predisposing developmental factors, such as childhood trauma, may continue their impact on DTD by impairing resilience and problem-solving. These social determinants may alter both the risk of developing and the likelihood of maintaining/recovering from DTD (Holzel, Harter, Reese, & Kriston, 2011; Negele, Kaufhold, Kallenbach, & Leuzinger-Bohleber, 2015; Verhoeven et al., 2020).

Although a single episode of depression that is responsive to treatment is unlikely to ‘scar’ personality (Shea et al., 1996), years of unremitting depression – with or without adequate treatment – may alter the clinical course and have taxonomic implications for DTD. Persons with DTD may develop behavioral and

thought patterns that exacerbate their negative self-valuation and pessimism. In essence, the depression builds on itself, intensifying self-defeating thought patterns and worsening the condition. Such ‘secondary’ changes in morale, self-efficacy and perceived resilience or grit may provide therapeutic opportunities (Thase & Howland, 1994; Young, 2018), as in helping patients with DTD to identify psychological ‘negative feedback loops’ and trigger the use of specific psychotherapeutic strategies to mitigate their impact (Eisendrath et al., 2015; Lynch et al., 2020).

The years of hopelessness associated with DTD may alter its course and have taxonomic implications. Persons with DTD may develop behavioral and thought patterns that exacerbate their negative self-valuation and pessimism. In essence, the depression builds on itself, intensifying self-defeating thought patterns and worsening the condition. This secondary demoralization may provide therapeutic opportunities. For example, identifying psychological ‘negative feedback loops’ could trigger the use of specific psychotherapeutic interventions to mitigate their impact.

The consideration of a DTD taxonomy raises concern about whether the psychology and biology of DTD evolve over time, as appears true in multiple medical conditions (e.g. congestive heart failure, atrial fibrillation, cancers). For many DTD patients, treatments effective earlier in their illness subsequently lose benefit, suggesting a developmental change in key neurobiological substrates (Katz, 2011). This observation also leads to a related and worrisome consideration: the possibility that exposure to ineffective or partially effective treatment induces neurobiological change such that treatment responsiveness diminishes and the depression becomes more difficult to treat (Andrews & Amsterdam, 2020; Andrews, Kornstein, Halberstadt, Gardner, & Neale, 2011; Fava, 2003a; Sackeim, 2016). In other words, treatment resistance could beget more profound treatment resistance and greater

chronicity. This possibility illustrates one of the many complex research challenges posed by DTD in developing a taxonomy.

Assessment of antidepressant treatment history

Knowledge about the nature and results of prior depression treatments is crucial to informing a taxonomy. This information (e.g. dose, duration, adherence, outcome) will be key in distinguishing among patients who do not receive 'adequate' therapeutic trials due to intolerance, those who show only minimal acute benefit despite 'adequate' treatment, and those who have greater acute responsivity but cannot sustain the benefit. At a practical level, such knowledge could also inform investigators as to which patients have already benefited or not from the intervention under study. In theory, a national electronic health record (EHR) would be optimal in providing universal and uniform data collection (Fife et al., 2017; Gronemann, Jorgensen, Nordentoft, Andersen, & Osler, 2020), but the USA is decades from that possibility. An alternative approach of melding different EHRs to create a continuous treatment narrative is feasible in principle, but a major challenge and expense.

To overcome these challenges, the field has developed several tools to retrospectively gather and evaluate treatment history, including the Antidepressant Treatment History Form (Sackeim, 2001; Sackeim et al., 2019), Maudsley Staging Method (Fekadu et al., 2009), and Massachusetts General Hospital-Antidepressant Treatment Questionnaire (Chandler, Iosifescu, Pollack, Targum, & Fava, 2010), among others (Gaynes et al., 2018). However, obtaining the requisite information is labor-intensive given the fractured nature of our health care systems, and the fact that past providers, pharmacies, medical facilities, patients, families, and caregivers may have differing and useful information that must be integrated. The rigor in establishing prior history, and thus the quality of the information obtained, likely differs among studies, which reduces consistency in findings. Furthermore, these tools have not been compared, and extensive validation studies have not been undertaken. Thus, while the assessment of treatment resistance is an integral part of DTD characterization and has shown predictive power regarding responsivity to subsequent treatment (Heijnen et al., 2010; Lisanby et al., 2009; Rush et al., 2006b) and relapse potential (Prudic et al., 2013; Sackeim et al., 1990), the assessment tools require further development.

Selecting, acquiring, and interpreting outcomes in DTD

Selecting among DTD outcomes

Until now, the primary outcome metric for evaluating MDE interventions has been depressive symptom severity, assessed over 6–12 week acute trials, and quantified either by a change in scores on a clinician-rated scale or by the proportion of participants with a clinically meaningful benefit specified categorically (e.g. remission, response, partial response) compared to control conditions. However, depressive symptoms may not necessarily be the most critical outcome. Some modest, but valuable, degree of symptom control has often already been achieved, and further meaningful symptomatic reduction is not expected, given the history of responsivity to prior treatments. Aside from symptom control, DTD treatments/interventions may aim to improve other factors, such as managing concurrent psychiatric and general medical conditions, minimizing treatment burden, enhancing daily

function, QoL, or overall mental and physical wellness, mitigating symptomatic worsening, or otherwise reducing mood instability (Fournier, DeRubeis, Amsterdam, Shelton, & Hollon, 2015; McAllister-Williams et al., 2020; Rush & Thase, 2018; Rush et al., 2019) (Fig. 2). From the care system-resource management perspective, cost efficiency is also an important DTD outcome, given its chronicity (Ross, Zivin, & Maixner, 2018).

To illustrate the potential importance of these holistic outcomes, patients who receive ECT are particularly characterized by their level of baseline impairment in social and vocational function and QoL. These deficits are typically as or more critical than symptom severity in leading to ECT referral, and typically resolve fully only with sustained remission following ECT (McCall, Prudic, Olfson, & Sackeim, 2006; McCall et al., 2013, 2017). Nevertheless, ECT trials have focused exclusively on the level of depressive symptom reduction as the primary outcome.

Choice of primary and secondary outcomes

A wide range of outcomes (with or without symptoms) can be targets for DTD interventions (Fig. 2). Researchers face the challenge of prioritizing amongst them. A single intervention may be aimed at one or more targets (e.g. symptom control and daily function). However, these various potential therapeutic outcomes may manifest at different times (e.g. symptom control may precede functional improvement by weeks or months) (Hofmann, Curtiss, Carpenter, & Kind, 2017; Paykel, 2002). There may also be trade-offs between achieving one goal (e.g. minimizing treatment burden) and optimizing function (Fournier et al., 2015).

Selecting a primary outcome would be far simpler had we an analog in DTD of the hemoglobin (Hgb) A1C measure in diabetes. This measure reflects the average glucose level in the bloodstream over the prior 2–3 months (thereby taking into account complex effects of multiple determinants). Hgb A1C is a strong indicator of disease process/control, with established benchmarks for normal, moderately severe, and severe dysregulation. It informs management decisions, and assessment of disease outcomes and complications (Sherwani, Khan, Ekzhaimy, Masood, & Sakharkar, 2016). Unfortunately, DTD is likely more heterogeneous with respect to etiology and pathophysiological processes, treatment responsiveness, and other factors, so it is unlikely that a single measure like Hgb A1C will emerge. Such a measure may be achievable within a specific DTD subgroup.

Symptom control in DTD refers to control of core criterion depressive or manic symptoms, as well as associated symptoms which can impact QoL, day-to-day functioning, and relapse risk. Such symptoms may include insomnia, pain, anxiety, irritability, cognitive impairment, and substance misuse. Which DTD outcome measures are chosen depends on the question(s) being addressed and the specific DTD subgroup under study. A reduction in symptomatic variability (waxing and waning) may be especially important in the management of the difficult-to-treat bipolar disorder, while addressing treatment burden and adherence may be key outcomes in another subgroup. We recognize that bipolar depressions are also often difficult-to-treat and the research and clinical challenges posed by bipolar DTD deserve separate in-depth discussion.

To enhance day-to-day function and QoL in DTD, optimizing control of medical and psychiatric co-morbidities may also be prioritized. This imposing number of potential primary and secondary outcomes in DTD presents challenges in selecting assessment domains and specific measures, and in determining

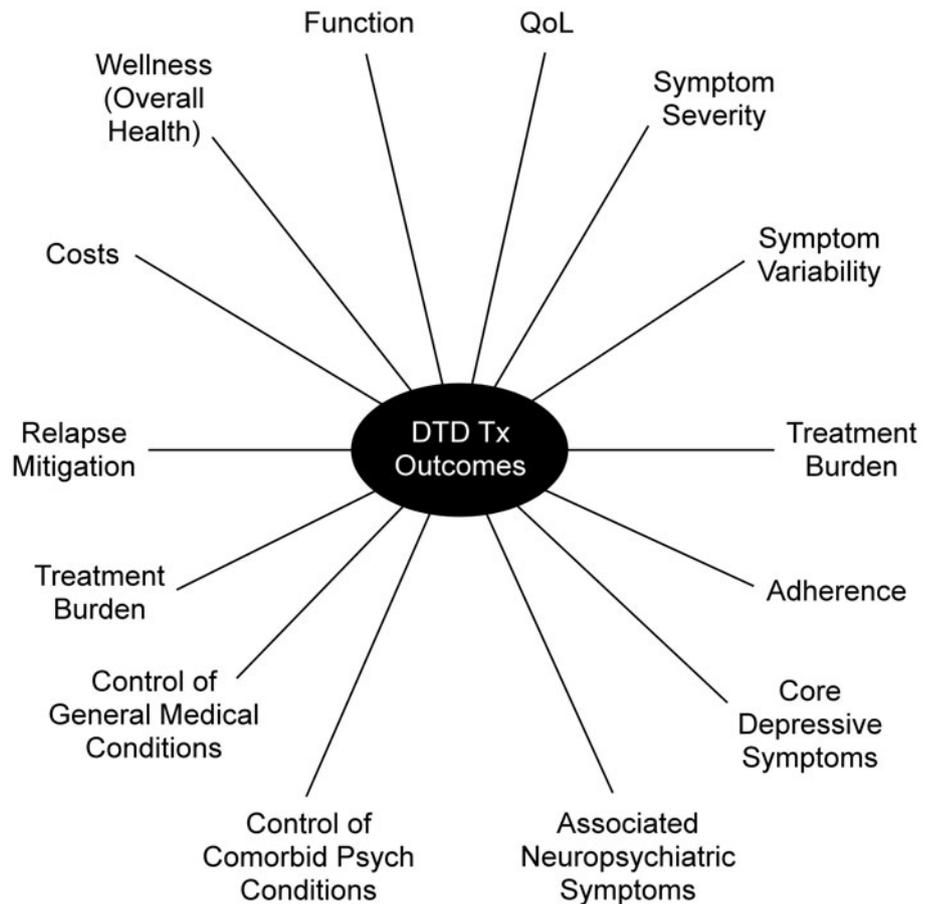


Fig. 2. Clinically important outcomes for DTD intervention research. Psych = Psychiatric; Tx = Treatment.

assessment frequency. Therefore, it is essential that studies are based on specific hypotheses that determine the outcome measures chosen.

Are multi-dimensional or composite outcomes needed for DTD?

The multifaceted outcomes in DTD raise the possibility of approximating a 'Hgb A1C-like' outcome metric by forming a composite or multidimensional outcome measure from the assessment of several diverse outcomes. Composite outcomes typically combine various aspects of a particular single construct, such as speed and extent of symptom change, or likelihood and persistence of such change. Multidimensional outcomes assess diverse domains that cannot be easily reflected in a single construct, such as the benefits and costs of an intervention (Schwartz & Patrick, 2014).

Figure 3 illustrates an attempt at evaluating outcomes using a multidimensional approach (Bech, 2009; Bech, Fava, Trivedi, Wisniewski, & Rush, 2012). Bech et al. (2012) applied the 'pharmacopsychometric triangle' which includes three dimensions that, when presented separately, enable clinicians and patients to see the trade-off between benefits and side effect burden (Bech, 2009). This approach meaningfully differentiated between buspirone and bupropion as adjunctive agents in depressed outpatients – a difference that was not observed when comparing effects on depressive symptoms alone (Trivedi et al., 2006).

Composite or combinatorial outcomes may have particular prognostic value in depression, whether or not specific to DTD.

The utility of combinatorial measures was illustrated by Cohen, Greenberg, and IsHak (2013). In a retrospective analysis of the STAR*D database, they found that three measures (symptom severity, function, and QoL) combined into a 'burden of illness scale' was better at predicting time to relapse following successful acute-phase antidepressant treatment than any element alone.

Regardless of whether standard scales, new composite, or new multidimensional measures are adopted as outcome metrics for DTD intervention studies, we must be able to translate changes in such scores into clinically useful categories that facilitate clinical decision-making and are meaningful to patients, care systems, and clinicians. This aim is analogous to the outcome categories of partial response, response, and remission, based on symptom change used in traditional acute antidepressant trials. For DTD patients, for whom various interventions with diverse aims (e.g. function, QoL, treatment burden, etc.) may be attempted, stakeholders will want to know what incremental changes are likely to be achieved in which groups of patients. Such a multidimensional/combinatorial model for DTD will require consensus as to what represents positive clinical outcomes.

How often and when should outcomes be obtained?

Typically, longitudinal outcome data address three aspects of change: (1) average change in the population; (2) individual differences in the degree of change; and (3) individual within-person change over different temporal intervals (Hofer, Thurvaldsson, & Piccinin, 2012). Therefore, the primary and secondary aims of a

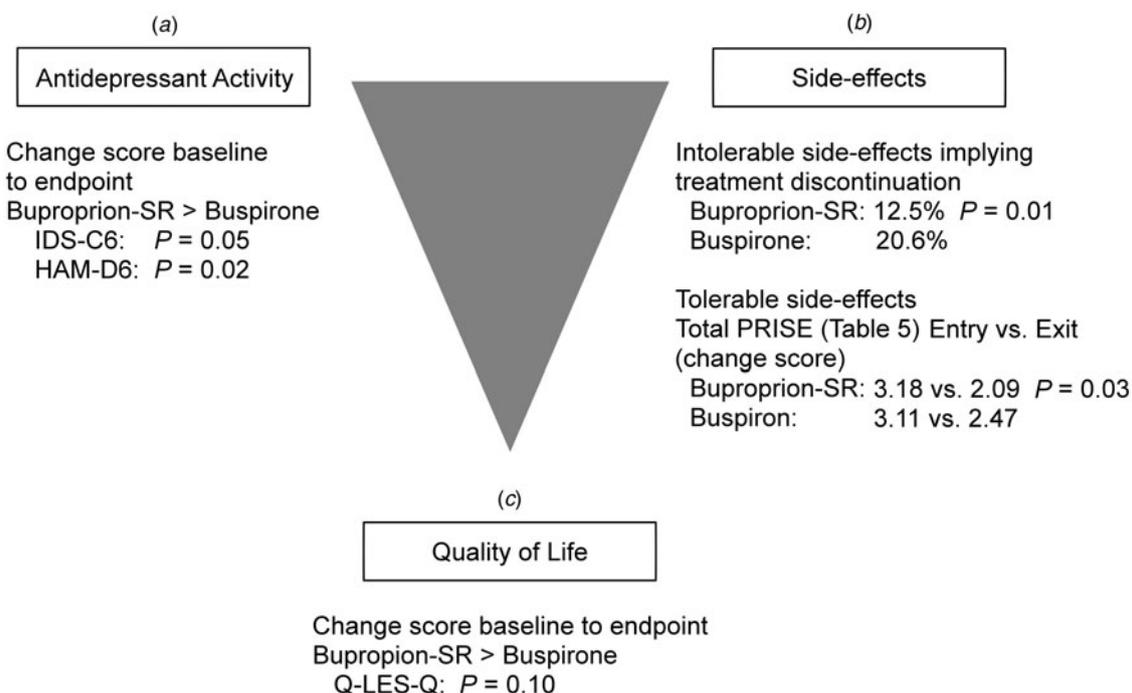


Fig. 3. Application of the pharmaco-psychometric triangle.

Note: Figure recreated from Bech et al. (2012). HAM-D6 = Hamilton Rating Scale for Depression 6-item subscale; IDS-C6 = Inventory of Depressive Symptomatology 6-item subscale – Clinician-rated; PRISE = Pragmatic-explanatory continuum indicator summary; Q-LES-Q = Quality of Life Enjoyment and Satisfaction Questionnaire; SR = Sustained release.

study must inform decisions regarding the frequency of obtaining outcomes of interest.

Some treatments are expected to differ radically in time to the onset of clinical effects. For example, symptomatic change with (es)ketamine may be observed within hours of the first infusion and, unless treatment is continued at regular intervals, usually wanes over the next 5–7 days (Zarate & Niciu, 2015). Other interventions, such as Vagus Nerve Stimulation, require substantially longer time frames (months to years) to fully manifest therapeutic benefit (Aaronson et al., 2017; Berry et al., 2013). Therefore, the frequency of repeated assessments should take into account variation in the outcome measure within and across individuals, the treatment being studied (Hofer et al., 2012), and the aim of treatment (e.g. acute symptom control or longer-term prophylaxis).

Different outcomes may need to be collected at different times because the time course for achieving these potentially important, but diverse, objectives is variable. For example, depressive symptom improvement often occurs before the full realization of improvement in function/QoL (McKnight & Kashdan, 2009). Given the temporal variability in symptom expression and the high rate of relapse following initial improvement, the durability of benefit is a key consideration in evaluating the therapeutic effects in DTD (Jelovac, Kolshus, & McLoughlin, 2013; Rush et al., 2006b; Sackeim et al., 2007) and considering the relative merits of alternative interventions (Cuijpers et al., 2013). Thus, the choices of outcomes, and when and how often they are measured, should be paramount when determining the duration of DTD treatment trials.

In addition, there can be substantial day-to-day variation in symptoms, function, and other outcome domains in DTD, either spontaneously or in reaction to environmental events. Furthermore, relatively small degrees of change consistently maintained may be

salutatory and even the main goal of treatment. However, detection of such limited degrees of change is highly contingent on the reliability of the baseline measures against which all subsequent assessments are compared. It is critical to allow adequate time (perhaps weeks) with potentially multiple averaged measures to establish a true baseline in DTD. Thus, the frequency of outcome assessments and the duration of the baseline period may be especially critical considerations when designing DTD intervention studies.

Which sources provide the most valid outcomes for DTD intervention research?

In typical acute-phase antidepressant trials with pharmacotherapy, psychotherapy, or ECT, there is a relatively moderate correlation between baseline self-report and clinician-ratings of symptom severity, often on the order of only 25% shared variance (Sayer et al., 1993; Uher et al., 2012). Of note, effect sizes for therapeutic interventions in MDEs are typically larger for clinician ratings than self-reports (Sayer et al., 1993). Such discrepancies may be both larger and more likely in DTD for which self-appraisal distortions or thinking biases caused by chronic illness or awareness of chronic shortcomings may further exaggerate negative self-assessment and impaired motivation. It is the common clinical experience for chronically depressed patients to be slower to recognize symptomatic improvement than clinicians who observe them at regular intervals. This suggests that the validity of self-report and the use of both self-reports and clinician ratings as combined outcomes in DTD deserve study. Older depressed persons show a consistent tendency to under self-report symptom severity relative to observer measures (Fiske, Wetherell, & Gatz, 2009; Gallo, Anthony, & Muthén, 1994). It is unknown

whether this aging effect is maintained, reversed, or accelerated in DTD, but given that DTD frequently persists well into old age, this factor must be considered. It would also be valuable to explore the utility of combining self- and observer-rated scales into a composite measure.

Early identification of mediators, moderators and predictors is essential

Due to major evidence gaps, clinicians treating individuals with DTD face several decision-making challenges, including (1) choosing among alternative treatments for a specific patient (identifying prescriptive predictors based on moderator analyses); (2) estimating the likelihood of benefit of a single treatment (based on baseline features, i.e. prognostic predictors); and (3) identifying mediators of treatment outcome [i.e. processes deemed to be essential to achieving the benefit (Kazdin, 2007)].

For example, consider a treatment aimed at enhancing psychological abilities to promote resilience-enhancing skills in everyday life, which in turn is expected to improve daily function or QoL. The measurement of resilience (the hypothetical mediator) is essential to determine whether the treatment outcomes achieved actually depend on enhanced resilience in those who benefit.

The identification of mediators, moderators, or prognostic predictors is especially important in DTD due to its etiological and treatment response heterogeneity. Any intervention will likely be useful for only a subset of DTD patients. Some outcomes may entail different mediators, moderators or prognostic predictors than others. From a cost-efficiency perspective, it is less expensive to select candidate measures as potential predictors and mediators early in the conduct of trials, whether observational or randomized, given the overall cost of conducting the trial (Trivedi *et al.*, 2016; Uher *et al.*, 2009). The identification of any mediator, moderator, or prognostic predictor could contribute to the taxonomy and to targeting the treatment to those most likely to benefit.

How should outcomes be collected?

To contain costs and gain granularity, we may need to consider different ways to acquire outcomes, especially to assess within-person change. This granularity will be critical when identifying mediators of change that gain traction at different time periods across different patients. For example, if reduced symptomatic variability is a mediator for improved QoL/daily function, this relationship may appear earlier in the course of some treatments than others. In addition, within a single treatment, some may show the response and thus reveal the mediator earlier than those who respond later. This variation across patients demands a level of granularity for outcome acquisition that can prove cost-prohibitive unless cleverly designed. Hypothetically, remote sampling via smartphones and passive collection of behavioral and physiological data could provide more precise and frequent outcome sampling. Indeed, intervention studies of DTD in the real world could be improved by delinking outcome assessment from treatment visits. Such visits are often scheduled based on how well the patient appears and other clinical considerations that, ideally, should be independent of the timing of outcome measurement. With the rise in telepsychiatry and data acquisition facilitated by natural language processing (NLP) and other forms of artificial intelligence (AI), remote real-world sampling of

outcomes (e.g. symptoms, function, or QoL) has become cost-effective and potentially may provide more valid outcomes.

Challenges in intervention trial design

This section discusses three trial design challenges presented by DTD: sample sourcing, trial execution, and intervention study designs that preserve causal inference.

Sample sourcing and eligibility criteria

DTD and TRD are heterogeneous in clinical presentation, course of illness, biology, treatment responsiveness, and other factors. Recruiting DTD intervention research participants from representative real-world treatment settings (as opposed to research clinics) would help ensure a representative sample which enhances the generalizability of findings and may reduce costs of recruitment, treatment delivery, and trial management costs. Large care systems, whether governmental (e.g. VA) or private commercial systems, can lower costs by providing much of the information on the prior treatment of mental health and general medical conditions.

Access to large numbers of representative participants helps to address the heterogeneity issue. Large numbers enable the identification of subgroups within the DTD patient population, especially those responsive or unresponsive to a particular intervention. This enables post-hoc secondary analyses to generate hypotheses for subsequent studies. Larger patient populations also increase the certainty of findings which, when obtained in real-world settings, enable rapid real-world implementation.

However, due to the heterogeneity of DTD and our current lack of taxonomy, there are risks in being too inclusive. Inclusion criteria must ensure that people with the targeted problem are included with few others, while the exclusion criteria might best be minimal to enhance generalizability. For example, a treatment targeting anhedonia in DTD might include depressed patients with sufficient baseline anhedonia to show the intended effect. In addition to the heterogeneity of DTD, potential efficacy-confounding aspects of DTD (e.g. the presence of a personality disorder, a background of substance misuse) are often not easily identifiable by standard review of diagnoses/medication exposure in large EHR databases. This is particularly challenging as concomitant personality disorder and substance addiction significantly impact depression outcomes (Davis *et al.*, 2006; Mulder, 2002). For the particular trial, the judicious assessment of factors that are suspected of affecting the outcome could be acquired on a full or subsample basis for secondary analyses.

Trial execution

To acquire large samples of representative DTD patients, point-of-care trials at multiple sites would be preferred with an emphasis on larger numbers of participants and, to contain costs, modest numbers of process and outcome measures (Fiore *et al.*, 2011; Shih, Turakhia, & Lai, 2015). Simple self-reported outcomes, perhaps collected via smartphones and/or at clinic visits (e.g. symptom burden, treatment burden/side effects, and function/QoL), can be both reliable and cost-efficient. Conveniently and frequently acquired global ratings may be as informative as measures derived from longer questionnaires, especially self-ratings on items or item response patterns that yield clinically meaningful outcome differences (Turkoz *et al.*, 2021).

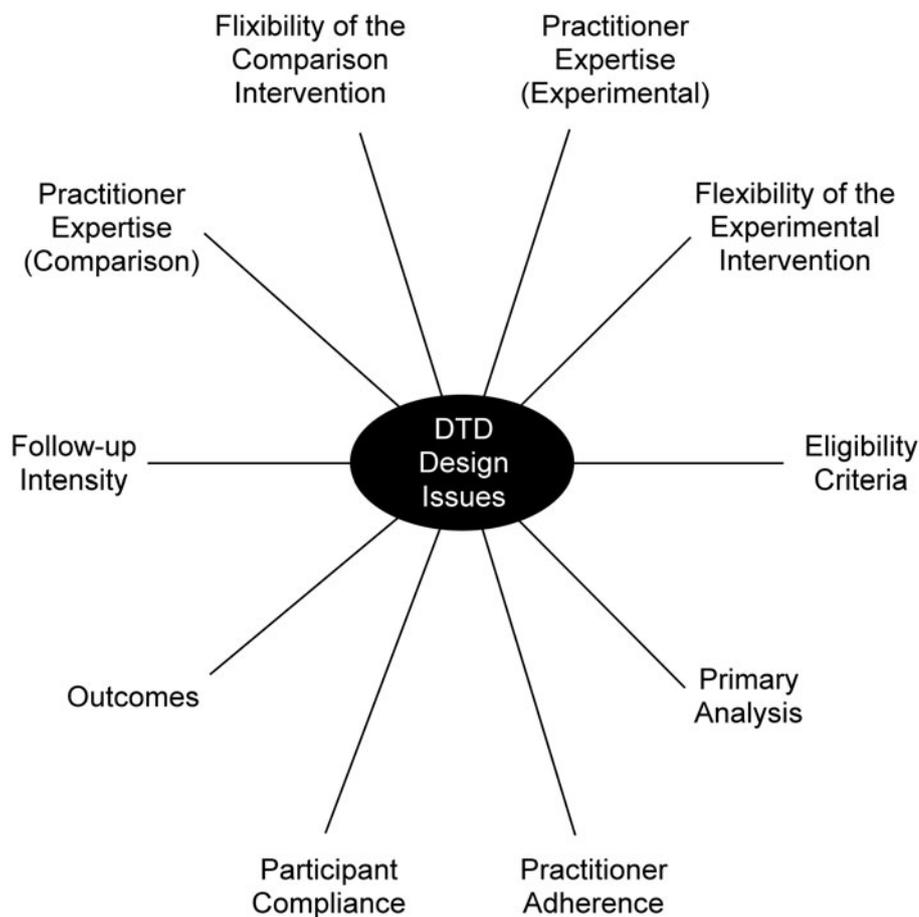


Fig. 4. Pragmatic-explanatory continuum indicator summary (PRECIS) wheel (Thorpe et al., 2009).

One strategy to identify the target sample for new interventions in DTD is the use of patient registries, in which initial testing of an intervention in an open trial identifies the subgroups for which the intervention is more or less effective (Aaronson et al., 2017; Sackeim et al., 2020). These registries would obtain patient-reported outcomes on a wide swath of DTD patients. Such information would form the basis for an evidence-based selection of a limited number of inclusion and exclusion criteria for use in subsequent randomized trials. This sort of registry would also identify individuals for whom the intervention causes untoward effects or no meaningful benefit, thereby enriching the study sample and reducing costs in the randomized controlled trial. In addition, preferred dosing and the expected trajectory of benefit would support a more time and cost-efficient randomized controlled trial design.

Study designs to optimize both generalizability and causal inference (hybrid trials)

A study of representative DTD patients treated in real-world settings raises a range of trial design considerations. Thorpe et al. (2009) identified 10 treatment trial design parameters that likely affect outcomes, and which vary considerably between highly controlled efficacy (explanatory) studies and effectiveness (pragmatic) studies. They include practitioner expertise in the interventions, eligibility criteria, follow-up frequency, outcomes, patient compliance, practitioner adherence, and flexibility of the experimental and comparison interventions (Fig. 4). These parameters are worth considering since variable delivery of the intervention

(s) in relation to these parameters has untoward consequences: (1) a very large sample may be needed to detect treatment effects; (2) ability to detect a signal (intervention difference) is reduced; and (3) if a difference is found between the interventions, the cause may be due to one or more of these parameters differentiating the two groups rather than differences in intrinsic therapeutic properties. Controlling these parameters in trial design can increase the certainty of inferring that the between-intervention differences (or non-differences) are due to the interventions and not an artifact of variability in delivery or selection bias.

This strategy enables causal attributions to be made when between-group differences are found in specified comparisons. For example, STAR*D engaged representative patients and practitioners while controlling the delivery of treatment using a measurement-based care guidance approach (Rush et al., 2004). This helped ensure that when treatment was unsuccessful, it was the treatment that failed and not its delivery. To make causal inferences from real-world studies, investigators should consider which parameters should be controlled, which will be assessed, and consequently, what kinds of causal inference can be made (Fig. 4).

Conclusions

In recent years, our expectations about the clinical utility of anti-depressant medications and psychotherapies in depressive and related mood disorders have been lowered, particularly for individuals who have a history of nonresponse to several standard

forms of therapy (Cuijpers, Karyotaki, Reijnders, & Ebert, 2019; Penn & Tracy, 2012). This expectational shift became apparent as intervention research moved to real-world patients with broader inclusion and fewer exclusion criteria, and as treatment delivery became less controlled (Bauer *et al.*, 2009; Rush *et al.*, 2004; Uher *et al.*, 2012). It is now recognized that only about one-third of those who receive an initial course of antidepressant pharmacotherapy will experience a sustained remission. Furthermore, previous treatment failure decreases the likelihood of achieving acute remission at the end of subsequent short-term medication trials, while also increasing the likelihood of relapse if remission is achieved (Conway *et al.*, 2017; Rush *et al.*, 2006b; Sackeim, 2016).

Despite the profound personal, familial, and societal costs of DTD, patients with DTD are often excluded from studies of therapeutic interventions and neurobiology. Even trials that specifically address therapeutics in TRD often cap the number of failed prior treatment trials or the duration of the current episode precisely because greater treatment resistance or chronicity is expected to negatively impact therapeutic outcomes, limiting detection of a therapeutic signal (O'Reardon *et al.*, 2007; Rush *et al.*, 2005). DTD patients are also excluded from research because their high level of medical and psychiatric co-morbidities, suicidality, or functional impairment may present safety concerns, complicate identification of therapeutic signals, or impose practical impediments to research participation.

Other branches of medicine have identified subgroups that do not benefit sufficiently from standard therapeutics and have developed methods to specifically study these populations and identify novel treatments or interventions that mitigate aspects of the clinical presentation. For example, consensus guidelines in epilepsy recommend consideration of surgical or neuromodulatory interventions following two unsuccessful trials of anticonvulsant medications (Brodie *et al.*, 2012; Callaghan *et al.*, 2007; Jette *et al.*, 2014; Kwan & Brodie, 2010). The European League Against Rheumatism defines difficult-to-treat rheumatoid arthritis (D2T RA) as persistent signs and/or symptoms despite a trial of two or more biologic or targeted synthetic disease-modifying anti-rheumatic drugs with different mechanisms of action, signs suggestive of active/progressive disease, and disease management viewed as problematic by the clinician and/or patient. Multiple differences have been identified in clinical presentation, treatment burden, and co-morbidities between D2T RA and comparison patients, and empirically defined subgroups within D2T RA have been proposed (Roodenrijs *et al.*, 2021). In contrast, we lack information on when in the course of multiple treatment trials DTD should be declared, with altered expectations regarding prognosis and the types of interventions considered. Indeed, since DTD has not been an object of study, we lack fundamental information on the nature and size of this population, demographic and clinical characteristics, optimal management, and long-term outcomes.

In conclusion, the traditional research methods used to develop and test therapeutic interventions in treatment-responsive or treatment-naïve mood disorder populations are often inapplicable in DTD. Indeed, our understanding of the natural history of mood disorders, the phases of illness, and the phases of treatment (e.g. continuation or maintenance regimens) may not readily apply to DTD (Frank *et al.*, 1991; Rush *et al.*, 2006a). Nonetheless, many individuals do not achieve sustained remission despite multiple well-delivered treatments. These individuals consume a disproportionate share of health

resources, while experiencing a disproportionate degree of functional impairment and prolonged suffering. Intervention research is sorely needed and should be particularly useful if it takes into account the clinical research challenges posed by DTD.

Acknowledgements. We thank Jon Kilner for his editorial assistance with manuscript preparation.

Financial support. The preparation of this manuscript was supported in part by LivaNova PLC. The funder had no role in determining its content.

Conflict of interest. Dr Rush has received consulting fees from Compass Inc., Curbstone Consultant LLC, Emmes Corp., Evecxia Therapeutics, Inc., Holmusk, Johnson and Johnson (Janssen), LivaNova PLC, Neurocrine Biosciences Inc., Otsuka-US, Sunovion; speaking fees from LivaNova, Johnson and Johnson (Janssen); and royalties from Guilford Press and the University of Texas Southwestern Medical Center, Dallas, TX (for the Inventory of Depressive Symptoms and its derivatives). He is also named co-inventor on two patents: U.S. Patent No. 7795 033: Methods to Predict the Outcome of Treatment with Antidepressant Medication, Inventors: McMahon FJ, Laje G, Manji H, Rush AJ, Paddock S, Wilson AS; and U.S. Patent No. 7 906 283: Methods to Identify Patients at Risk of Developing Adverse Events During Treatment with Antidepressant Medication, Inventors: McMahon FJ, Laje G, Manji H, Rush AJ, Paddock S. Dr Sackeim serves as a scientific adviser to Cerebral Therapeutics Inc., LivaNova PLC, MECTA Corporation, Neuroief Ltd., Neuronetics Inc, and Parow Entheobiosciences LLC. He receives honoraria and royalties from Elsevier Inc. and Oxford University Press. He is the inventor on non-remunerative US patents for Focal Electrically Administered Seizure Therapy (FEAST), titration in the current domain in ECT, and the adjustment of current in ECT devices, each held by the MECTA Corporation. He is also the originator of magnetic seizure therapy (MST). Dr Conway has received research support from Bristol-Myers Squibb, the Stanley Medical Research Institute, the National Institute of Mental Health, NeoSync, LivaNova PLC, the Taylor Family Institute for Innovative Psychiatric Research, the American Foundation for Suicide Prevention, Assure Health Inc., August Busch IV Foundation, and Barnes-Jewish Hospital Foundation. He is a part-time employee at the John Cochran VA Medical Center in St. Louis. Dr Bunker is a former employee and a current consultant of LivaNova USA PLC. Dr Hollon has no disclosures. Dr Demyttenaere has received an honorarium for attending advisory boards, acting as a consultant or being a member of the speaker bureau for Boehringer-Ingelheim, Gedeon-Richter, Johnson and Johnson, LivaNova, Lundbeck, Pfizer and Recordati. Dr Young has received payment for lectures and advisory boards for the following companies: Astrazenaca, Eli Lilly, Lundbeck, Sunovion, Servier, LivaNova, Janssen, Allegan, Bionomics, Sumitomo Dainippon Pharma, COMPASS. He is a consultant to Johnson & Johnson and LivaNova. He has received honoraria for attending advisory boards and presenting talks at meetings organized by LivaNova. He is a Principal Investigator on studies funded by LivaNova; Janssen, COMPASS and Chief Investigator on a study funded by Novartis. He does not hold shares in pharmaceutical companies. Dr Aaronson is a consultant to Neuronetics, LivaNova, Janssen, Sage Therapeutics and Genomind. He also receives research support from Compass Pathways and Neuronetics. Dr Dibué is an employee of LivaNova PLC, a manufacturer of vagus nerve stimulators and holds stock options. Dr Thase reports the following relationships: Advisory/Consultant: Acadia, Inc., Akili, Inc., Alkermes PLC, Allergan, Inc., Clexio, H. Lundbeck, A/S, Jazz Pharmaceuticals, Johnson & Johnson (Janssen), Merck & Company, Inc., Otsuka Pharmaceutical Company, Ltd., Pfizer, Inc., and Seelos. Grant Support: Acadia, Inc., Allergan, Inc., AssureRx (now Myriad), Axsome Therapeutics Inc., Intracellular, Inc., Johnson & Johnson (Janssen), Otsuka Pharmaceutical Company, Ltd., Patient-Centered Outcomes Research Institute (PCORI), Takeda. Royalties: American Psychiatric Press, Guilford Publications, W.W. Norton & Company, Inc., and Wolters Kluwer. Spouse's Employment: Dr McAllister-Williams has received fees from American Center for Psychiatry & Neurology, United Arab Emirates, British Association for Psychopharmacology, European College of Neuropsychopharmacology, International Society for Affective Disorders,

Janssen, LivaNova, Lundbeck, My Tomorrows, OCM Comunicazione s.n.c., Pfizer, Qatar International Mental Health Conference, Sunovion, Syntropharma, UK Medical Research Council and Wiley; grant support from National Institute for Health Research Efficacy and Mechanism Evaluation Panel and Health Technology Assessment Panel; and non-financial support from COMPASS Pathways and Magstim.

References

- Aaronson, S. T., Sears, P., Ruvuna, F., Bunker, M., Conway, C. R., Dougherty, D. D., ... Zajecka, J. M. (2017). A 5-year observational study of patients with treatment-resistant depression treated with vagus nerve stimulation or treatment as usual: Comparison of response, remission, and suicidality. *American Journal of Psychiatry*, *174*, 640–648. doi: 10.1176/appi.ajp.2017
- Agency for Healthcare Research and Policy (AHQR) (Sept 2011). *Nonpharmacologic interventions for treatment-resistant depression in adults (comparative effectiveness review number 33)*. Rockville, MD: AHQR.
- Amital, D., Fostick, L., Silberman, A., Calati, R., Spindelegger, C., Serretti, A., ... Zohar, J. (2013). Physical co-morbidity among treatment-resistant vs. treatment responsive patients with major depressive disorder. *European Neuropsychopharmacology*, *23*, 895–901. doi: 10.1016/j.euroneuro.2012.09.002
- Amos, T. B., Tandon, N., Lefebvre, P., Pilon, D., Kamstra, R. L., Pivneva, I., & Greenberg, P. E. (2018). Direct and indirect cost burden and change of employment status in treatment-resistant depression: A matched cohort study using a US commercial claims database. *Journal of Clinical Psychiatry*, *79*(2), 17m11725. doi: 10.4088/JCP.17m11725
- Andrews, P., & Amsterdam, J. (2020). A hormetic approach to understanding antidepressant effectiveness and the development of antidepressant tolerance – A conceptual view. *Psychiatria Polska*, *54*, 1067–1090. doi: 10.12740/PP/120084
- Andrews, P., Kornstein, S. G., Halberstadt, L. J., Gardner, C. O., & Neale, M. C. (2011). Blue again: Perturbational effects of antidepressants suggest monoaminergic homeostasis in major depression. *Frontiers in Psychology*, *2*, 159. doi: 10.3389/fpsyg.2011.00159
- Arnaldi, G., Angeli, A., Atkinson, A. B., Bertagna, X., Cavagnini, F., Chrousos, G. P., ... Boscaro, M. (2003). Diagnosis and complications of Cushing's syndrome: A consensus statement. *The Journal of Clinical Endocrinology & Metabolism*, *88*, 5593–5602. doi: 10.1210/jc.2003-030871
- Bauer, M., Pfennig, A., Linden, M., Smolka, M. N., Neu, P., & Adli, M. (2009). Efficacy of an algorithm-guided treatment compared with treatment as usual: A randomized, controlled study of inpatients with depression. *Journal of Clinical Psychopharmacology*, *29*, 327–333. doi: 10.1097/JCP.0b013e3181ac4839
- Bech, P. (2009). Applied psychometrics in clinical psychiatry: The pharmacopsychometric triangle. *Acta Psychiatrica Scandinavica*, *120*, 400–409. doi: 10.1111/j.1600-0447.2009.01445.x
- Bech, P., Fava, M., Trivedi, M. H., Wisniewski, S. R., & Rush, A. J. (2012). Outcomes on the pharmacopsychometric triangle in bupropion-SR vs. buspirone augmentation of citalopram in the STAR*D trial. *Acta Psychiatrica Scandinavica*, *125*, 342–348. doi: 10.1111/j.1600-0447.2011.01791.x
- Benson, C., Szukis, H., Sheehan, J. J., Alphas, L., & Yuce, H. (2020). An evaluation of the clinical and economic burden among older adult medicare-covered beneficiaries with treatment-resistant depression. *The American Journal of Geriatric Psychiatry*, *28*, 350–362. doi: 10.1016/j.jagp.2019.10.012
- Berlim, M. T., & Turecki, G. (2007a). Definition, assessment, and staging of treatment-resistant refractory major depression: A review of current concepts and methods. *The Canadian Journal of Psychiatry*, *52*, 46–54. doi: 10.1177/070674370705200108
- Berlim, M. T., & Turecki, G. (2007b). What is the meaning of treatment-resistant/refractory major depression (TRD)? A systematic review of current randomized trials. *European Neuropsychopharmacology*, *17*, 696–707. doi: 10.1016/j.euroneuro.2007.03.009
- Berry, S. M., Broglio, K., Bunker, M., Jayewardene, A., Olin, B., & Rush, A. J. (2013). A patient-level meta-analysis of studies evaluating vagus nerve stimulation therapy for treatment-resistant depression. *Medical Devices (Auckl)*, *6*, 17–35. doi: 10.2147/MDER.S41017
- Brodie, M. J., Barry, S. J., Bamagous, G. A., Norrie, J. D., & Kwan, P. (2012). Patterns of treatment response in newly diagnosed epilepsy. *Neurology*, *78*, 1548–1554. doi: 10.1212/WNL.0b013e3182563b19
- Callaghan, B. C., Anand, K., Hesdorffer, D., Hauser, W. A., & French, J. A. (2007). Likelihood of seizure remission in an adult population with refractory epilepsy. *Annals of Neurology*, *62*, 382–389. doi: 10.1002/ana.21166
- Chandler, G. M., Iosifescu, D. V., Pollack, M. H., Targum, S. D., & Fava, M. (2010). RESEARCH: Validation of the Massachusetts general hospital anti-depressant treatment history questionnaire (ATRQ). *CNS Neuroscience & Therapeutics*, *16*, 322–325. doi: 10.1111/j.1755-5949.2009.00102.x
- Cohen, R. M., Greenberg, J. M., & IsHak, W. W. (2013). Incorporating multi-dimensional patient-reported outcomes of symptom severity, functioning, and quality of life in the individual burden of illness index for depression to measure treatment impact and recovery in MDD. *JAMA Psychiatry*, *70*, 343–350. doi: 10.1001/jamapsychiatry.2013.286
- Conway, C. R., George, M. S., & Sackeim, H. A. (2017). Towards an evidence-based, operational definition of treatment-resistant depression: When enough is enough. *JAMA Psychiatry*, *74*, 9–10. doi: 10.1001/jamapsychiatry.2016.2586
- Cuijpers, P., Berking, M., Andersson, G., Quigley, L., Kleiboer, A., & Dobson, K. S. (2013). A meta-analysis of cognitive-behavioural therapy for adult depression, alone and in comparison with other treatments. *The Canadian Journal of Psychiatry*, *58*, 376–385. doi: 10.1177/070674371305800702
- Cuijpers, P., Karyotaki, E., Reijnders, M., & Ebert, D. D. (2019). Was Eysenck right after all? A reassessment of the effects of psychotherapy for adult depression. *Epidemiology and Psychiatric Sciences*, *28*, 21–30. doi: 10.1017/S2045796018000057
- Davis, L. L., Frazier, E., Husain, M. M., Warden, D., Trivedi, M., Fava, M., ... Rush, A. J. (2006). Substance use disorder comorbidity in major depressive disorder: A confirmatory analysis of the STAR*D cohort. *The American Journal on Addictions*, *15*, 278–285. doi: 10.1080/10550490600754317
- DeRubeis, R. J., Zajecka, J., Shelton, R. C., Amsterdam, J. D., Fawcett, J., Xu, C., ... Hollon, S. D. (2020). Prevention of recurrence after recovery from a major depressive episode with antidepressant medication alone or in combination with cognitive-behavioral therapy: Phase 2 of a 2-phase randomized clinical trial. *JAMA Psychiatry*, *77*, 237–245. doi: 10.1001/jamapsychiatry.2019.3900
- Dunner, D. L., Rush, A. J., Russell, J. M., Burke, M., Woodard, S., Wingard, P., & Allen, J. (2006). Prospective, long-term, multicenter study of the naturalistic outcomes of patients with treatment-resistant depression. *Journal of Clinical Psychiatry*, *67*, 688–695. doi: 10.4088/jcp.v67n0501
- Duntas, L. H., & Maillis, A. (2013). Hypothyroidism and depression: Salient aspects of pathogenesis and management. *Minerva Endocrinologica*, *38*, 365–377.
- Eaton, W. W., Shao, H., Nestadt, G., Lee, H. B., Bienvenu, O. J., & Zandi, P. (2008). Population-based study of first onset and chronicity in major depressive disorder. *Archives of General Psychiatry*, *65*, 513–520. doi: 10.1001/archpsyc.65.5.513
- Eisendrath, S. J., Gillung, E., Delucchi, K., Mathalon, D. H., Yang, T. T., Satre, D. D., ... Wolkowitz, O. M. (2015). A preliminary study: Efficacy of mindfulness-based cognitive therapy versus sertraline as first-line treatments for major depressive disorder. *Mindfulness (N Y)*, *6*, 475–482. doi: 10.1007/s12671-014-0280-8
- Fava, G. A. (2003a). Can long-term treatment with antidepressant drugs worsen the course of depression? *Journal of Clinical Psychiatry*, *64*, 123–133. doi: 10.4088/jcp.v64n0204
- Fava, M. (2003b). Diagnosis and definition of treatment-resistant depression. *Biological Psychiatry*, *53*, 649–659. doi: 10.1016/s0006-3223(03)00231-2
- Fekadu, A., Wooderson, S., Donaldson, C., Markopoulou, K., Masterson, B., Poon, L., & Cleare, A. J. (2009). A multidimensional tool to quantify treatment resistance in depression: The Maudsley staging method. *Journal of Clinical Psychiatry*, *70*, 177–184. doi: 10.4088/jcp.08m04309
- Fife, D., Feng, Y., Wang, M. Y., Chang, C. J., Liu, C. Y., Juang, H. T., ... Wang, B. (2017). Epidemiology of pharmaceutically treated depression and treatment-resistant depression in Taiwan. *Psychiatry Research*, *252*, 277–283. doi: 10.1016/j.psychres.2017.03.006

- Fiore, L. D., Brophy, M., Ferguson, R. E., D'Avolio, L., Hermos, J. A., Lew, R. A., ... Lavori, P. W. (2011). A point-of-care clinical trial comparing insulin administered using a sliding scale versus a weight-based regimen. *Clinical Trials*, 8, 183–195. doi: 10.1177/1740774511398368
- Fiske, A., Wetherell, J. L., & Gatz, M. (2009). Depression in older adults. *Annual Review of Clinical Psychology*, 5, 363–389. doi: 10.1146/annurev.clinpsy.032408.153621
- Fournier, J. C., DeRubeis, R. J., Amsterdam, J., Shelton, R. C., & Hollon, S. D. (2015). Gains in employment status following antidepressant medication or cognitive therapy for depression. *The British Journal of Psychiatry*, 206, 332–338. doi: 10.1192/bjp.bp.113.133694
- Frank, E., Prien, R. F., Jarrett, R. B., Keller, M. B., Kupfer, D. J., Lavori, P. W., ... Weissman, M. M. (1991). Conceptualization and rationale for consensus definitions of terms in major depressive disorder. Remission, recovery, relapse, and recurrence. *Archives Of General Psychiatry*, 48, 851–855. doi: 10.1001/archpsyc.1991.01810330075011
- Gallo, J. J., Anthony, J. C., & Muthén, B. O. (1994). Age differences in the symptoms of depression: A latent trait analysis. *Journal of Gerontology*, 49, P251–P264. doi: 10.1093/geronj/49.6.p251
- Gaynes, B. N., Asher, G., Gartlehner, G., Hoffman, V., Green, J., Boland, E., ... Lohr, K. N. (2018). AHRQ Technology assessments. In *Definition of treatment-resistant depression in the medicare population* (pp. 18–35). Rockville, MD: Agency for Healthcare Research and Quality (US).
- Gaynes, B. N., Lux, L., Gartlehner, G., Asher, G., Forman-Hoffman, V., Green, J., ... Lohr, K. N. (2020). Defining treatment-resistant depression. *Depression and Anxiety*, 37, 134–145. doi: 10.1002/da.22968
- Greenberg, P., Corey-Lisle, P. K., Birnbaum, H., Marynchenko, M., & Claxton, A. (2004). Economic implications of treatment-resistant depression among employees. *Pharmacoeconomics*, 22, 363–373. doi: 10.2165/00019053-200422060-00003
- Gronemann, F. H., Jorgensen, M. B., Nordentoft, M., Andersen, P. K., & Osler, M. (2020). Socio-demographic and clinical risk factors of treatment-resistant depression: A Danish population-based cohort study. *Journal of Affective Disorders*, 261, 221–229. doi: 10.1016/j.jad.2019.10.005
- Hage, M. P., & Azar, S. T. (2012). The link between thyroid function and depression. *Journal of Thyroid Research*, 2012, 590648. doi: 10.1155/2012/590648
- Heijnen, W. T., Birkenhager, T. K., Wierdsma, A. I., & van den Broek, W. W. (2010). Antidepressant pharmacotherapy failure and response to subsequent electroconvulsive therapy: A meta-analysis. *Journal of Clinical Psychopharmacology*, 30, 616–619. doi: 10.1097/JCP.0b013e3181ee0f5f
- Hofer, S. M., Thurvaldsson, V., & Piccinin, A. M. (2012). Foundational issues of design and measurement in developmental research. In B. Laursen, T. D. Little & N. A. Card (Eds.), *Handbook of developmental research methods* (pp. 3–16). New York: Guilford Press.
- Hofmann, S. G., Curtiss, J., Carpenter, J. K., & Kind, S. (2017). Effect of treatments for depression on quality of life: A meta-analysis. *Cognitive Behaviour Therapy*, 46, 265–286. doi: 10.1080/16506073.2017.1304445
- Hollon, S. D., DeRubeis, R. J., Fawcett, J., Amsterdam, J. D., Shelton, R. C., Zajecka, J., ... Gallop, R. (2014). Effect of cognitive therapy with antidepressant medications vs antidepressants alone on the rate of recovery in major depressive disorder: A randomized clinical trial. *JAMA Psychiatry*, 71, 1157–1164. doi: 10.1001/jamapsychiatry.2014.1054
- Holzel, L., Harter, M., Reese, C., & Kriston, L. (2011). Risk factors for chronic depression--a systematic review. *Journal of Affective Disorders*, 129, 1–13. doi: 10.1016/j.jad.2010.03.025
- Huang, S. S., Chen, H. H., Wang, J., Chen, W. J., Chen, H. C., & Kuo, P. H. (2020). Investigation of early and lifetime clinical features and comorbidities for the risk of developing treatment-resistant depression in a 13-year nationwide cohort study. *BMC Psychiatry*, 20, 541. doi: 10.1186/s12888-020-02935-z
- Jaffe, D. H., Rive, B., & Denee, T. R. (2019). The humanistic and economic burden of treatment-resistant depression in Europe: A cross-sectional study. *BMC Psychiatry*, 19, 247. doi: 10.1186/s12888-019-2222-4
- Jelovac, A., Kolshus, E., & McLoughlin, D. M. (2013). Relapse following successful electroconvulsive therapy for major depression: A meta-analysis. *Neuropsychopharmacology* 38, 2467–2474. doi: 10.1038/npp.2013.149
- Jette, N., Reid, A. Y., & Wiebe, S. (2014). Surgical management of epilepsy. *Canadian Medical Association Journal*, 186, 997–1004. doi: 10.1503/cmaj.121291
- Katz, G. (2011). Tachyphylaxis/tolerance to antidepressive medications: A review. *The Israel Journal of Psychiatry and Related Sciences*, 48, 129–135.
- Kazdin, A. E. (2007). Mediators and mechanisms of change in psychotherapy research. *Annual Review of Clinical Psychology*, 3, 1–27. doi: 10.1146/annurev.clinpsy.3.022806.091432
- Kubitz, N., Mehra, M., Potluri, R. C., Garg, N., & Cossrow, N. (2013). Characterization of treatment-resistant depression episodes in a cohort of patients from a US commercial claims database. *PLoS One*, 8, e76882. doi: 10.1371/journal.pone.0076882
- Kwan, P., & Brodie, M. J. (2010). Definition of refractory epilepsy: Defining the indefinable? *The Lancet Neurology*, 9, 27–29. doi: 10.1016/S1474-4422(09)70304-7
- Lisanby, S. H., Husain, M. M., Rosenquist, P. B., Maixner, D., Gutierrez, R., Krystal, A., ... George, M. S. (2009). Daily left prefrontal repetitive transcranial magnetic stimulation in the acute treatment of major depression: Clinical predictors of outcome in a multisite, randomized controlled clinical trial. *Neuropsychopharmacology*, 34, 522–534. doi: 10.1038/npp.2008.118
- Lynch, T. R., Hempel, R. J., Whalley, B., Byford, S., Chamba, R., Clarke, P., ... Russell, I. T. (2020). Refractory depression-mechanisms and efficacy of radically open dialectical behaviour therapy (Reframed): Findings of a randomised trial on benefits and harms. *The British Journal of Psychiatry*, 216, 204–212. doi: 10.1192/bjp.2019.53
- McAllister-Williams, R. H., Arango, C., Blier, P., Demyttenaere, K., Falkai, P., Gorwood, P., ... Rush, A. J. (2020). The identification, assessment and management of difficult-to-treat depression: An international consensus statement. *Journal of Affective Disorders*, 267, 264–282. doi: 10.1016/j.jad.2020.02.023
- McCall, W. V., Lisanby, S. H., Rosenquist, P. B., Dooley, M., Husain, M. M., Knapp, R. G., ... Group, C. P. W. (2017). Effects of a right unilateral ultra-brief pulse electroconvulsive therapy course on health-related quality of life in elderly depressed patients. *Journal of Affective Disorders*, 209, 39–45. doi: 10.1016/j.jad.2016.11.003
- McCall, W. V., Prudic, J., Olfson, M., & Sackeim, H. (2006). Health-related quality of life following ECT in a large community sample. *Journal of Affective Disorders*, 90, 269–274. doi: 10.1016/j.jad.2005.12.002
- McCall, W. V., Reboussin, D., Prudic, J., Haskett, R. F., Isenberg, K., Olfson, M., ... Sackeim, H. A. (2013). Poor health-related quality of life prior to ECT in depressed patients normalizes with sustained remission after ECT. *Journal of Affective Disorders*, 147, 107–111. doi: 10.1016/j.jad.2012.10.018
- McKnight, P. E., & Kashdan, T. B. (2009). The importance of functional impairment to mental health outcomes: A case for reassessing our goals in depression treatment research. *Clinical Psychology Review*, 29, 243–259. doi: 10.1016/j.cpr.2009.01.005
- Mulder, R. T. (2002). Personality pathology and treatment outcome in major depression: A review. *American Journal of Psychiatry*, 159, 359–371. doi: 10.1176/appi.ajp.159.3.359
- Negele, A., Kaufhold, J., Kallenbach, L., & Leuzinger-Bohleber, M. (2015). Childhood trauma and its relation to chronic depression in adulthood. *Depression Research and Treatment*, 2015, 650804. doi: 10.1155/2015/650804
- Olchanski, N., McInnis Myers, M., Halseth, M., Cyr, P. L., Bockstedt, L., Goss, T. F., & Howland, R. H. (2013). The economic burden of treatment-resistant depression. *Clinical Therapeutics*, 35, 512–522. doi: 10.1016/j.clinthera.2012.09.001
- Olfson, M., Amos, T. B., Benson, C., McRae, J., & Marcus, S. C. (2018). Prospective service use and health care costs of Medicaid beneficiaries with treatment-resistant depression. *Journal of Managed Care & Specialty Pharmacy*, 24, 226–236. doi: 10.18553/jmcp.2018.24.3.226
- O'Reardon, J. P., Solvason, H. B., Janicak, P. G., Sampson, S., Isenberg, K. E., Nahas, Z., ... Sackeim, H. A. (2007). Efficacy and safety of transcranial magnetic stimulation in the acute treatment of major depression: A multisite randomized controlled trial. *Biological Psychiatry*, 62, 1208–1216. doi: 10.1016/j.biopsych.2007.01.018

- Paykel, E. S. (2002). Achieving gains beyond response. *Acta Psychiatrica Scandinavica*, 106, 12–17. doi: 10.1034/j.1600-0447.106.s415.3.x
- Penn, E., & Tracy, D. K. (2012). The drugs don't work? Antidepressants and the current and future pharmacological management of depression. *Therapeutic Advances in Psychopharmacology*, 2, 179–188. doi: 10.1177/2045125312445469
- Pivonello, R., Isidori, A. M., De Martino, M. C., Newell-Price, J., Biller, B. M., & Colao, A. (2016). Complications of Cushing's syndrome: State of the art. *The Lancet Diabetes & Endocrinology*, 4, 611–629. doi: 10.1016/S2213-8587(16)00086-3
- Prudic, J., Haskett, R. F., McCall, W. V., Isenberg, K., Cooper, T., Rosenquist, P. B., ... Sackeim, H. A. (2013). Pharmacological strategies in the prevention of relapse after electroconvulsive therapy. *The Journal of ECT*, 29, 3–12. doi: 10.1097/YCT.0b013e31826ea8c4
- Prudic, J., Haskett, R. F., Mulsant, B., Malone, K. M., Pettinati, H. M., Stephens, S., ... Sackeim, H. A. (1996). Resistance to antidepressant medications and short-term clinical response to ECT. *American Journal of Psychiatry*, 153, 985–992. doi: 10.1176/ajp.153.8.985
- Rasmussen, K. G., Mueller, M., Rummans, T. A., Husain, M. M., Petrides, G., Knapp, R. G., ... Kellner, C. H. (2009). Is baseline medication resistance associated with potential for relapse after successful remission of a depressive episode with ECT? Data from the consortium for research on electroconvulsive therapy (CORE). *Journal of Clinical Psychiatry* 70, 232–237. doi: 10.4088/jcp.08m04092
- Roodenrijs, N. M. T., van der Goes, M. C., Welsing, P. M. J., Tekstra, J., Lafeber, F. P. J. G., Jacobs, J. W. G., & van Laar, J. M. (2021). Difficult-to-treat rheumatoid arthritis: Contributing factors and burden of disease. *Rheumatology*, 60, 3778–3788. doi: 10.1093/rheumatology/keaa860
- Ross, E. L., Zivin, K., & Maixner, D. F. (2018). Cost-effectiveness of electroconvulsive therapy vs pharmacotherapy/psychotherapy for treatment-resistant depression in the United States. *JAMA Psychiatry*, 75, 713–722. doi: 10.1001/jamapsychiatry.2018.0768
- Rush, A. J., Aaronson, S. T., & Demyttenaere, K. (2019). Difficult-to-treat depression: A clinical and research roadmap for when remission is elusive. *Australian & New Zealand Journal of Psychiatry*, 53, 109–118. doi: 10.1177/0004867418808585
- Rush, A. J., Fava, M., Wisniewski, S. R., Lavori, P. W., Trivedi, M. H., Sackeim, H. A., ... Group, S. D. I. (2004). Sequenced treatment alternatives to relieve depression (STAR*D): Rationale and design. *Controlled Clinical Trials*, 25, 119–142. doi: 10.1016/s0197-2456(03)00112-0
- Rush, A. J., Kraemer, H. C., Sackeim, H. A., Fava, M., Trivedi, M. H., Frank, E., ... Schatzberg, A. F. (2006a). Report by the ACNP task force on response and remission in major depressive disorder. *Neuropsychopharmacology*, 31, 1841–1853. doi: 10.1038/sj.npp.1301131
- Rush, A. J., Marangell, L. B., Sackeim, H. A., George, M. S., Brannan, S. K., Davis, S. M., ... Cooke, R. G. (2005). Vagus nerve stimulation for treatment-resistant depression: A randomized, controlled acute phase trial. *Biological Psychiatry*, 58, 347–354. doi: 10.1016/j.biopsych.2005.05.025
- Rush, A. J., & Thase, M. E. (2018). Improving depression outcome by patient-centered medical management. *American Journal of Psychiatry*, 175, 1187–1198. doi: 10.1176/appi.focus.18207
- Rush, A. J., Trivedi, M. H., Wisniewski, S. R., Nierenberg, A. A., Stewart, J. W., Warden, D., ... Fava, M. (2006b). Acute and longer-term outcomes in depressed outpatients requiring one or several treatment steps: A STAR*D report. *American Journal of Psychiatry*, 163, 1905–1917. doi: 10.1176/ajp.2006.163.11.1905
- Rush, A. J., Wisniewski, S. R., Zisook, S., Fava, M., Sung, S. C., Haley, C. L., ... Hollon, S. D. (2012). Is prior course of illness relevant to acute or longer-term outcomes in depressed out-patients? A STAR*D report. *Psychological Medicine*, 42, 1131–1149. doi: 10.1017/S0033291711002170
- Sackeim, H. A. (2001). The definition and meaning of treatment-resistant depression. *Journal of Clinical Psychiatry*, 62(Suppl. 16), 10–17.
- Sackeim, H. A. (2016). Acute continuation and maintenance treatment of major depressive episodes with transcranial magnetic stimulation. *Brain Stimulation*, 9, 313–319. doi: 10.1016/j.brs.2016.03.006
- Sackeim, H. A. (2021). Staging and combining brain stimulation interventions: Vagus nerve stimulation and electroconvulsive therapy. *Journal of ECT*, 37, 80–83. doi: 10.1097/YCT.0000000000000745
- Sackeim, H. A., Aaronson, S. T., Bunker, M. T., Conway, C. R., Demitrack, M. A., George, M. S., ... Rush, A. J. (2019). The assessment of resistance to antidepressant treatment: Rationale for the antidepressant treatment history form: Short form (ATHF-SF). *Journal of Psychiatric Research*, 113, 125–136. doi: 10.1016/j.jpsychires.2019.03.021
- Sackeim, H. A., Aaronson, S. T., Carpenter, L. L., Hutton, T. M., Mina, M., Pages, K., ... West, W. S. (2020). Clinical outcomes in a large registry of patients with major depressive disorder treated with transcranial magnetic stimulation. *Journal of Affective Disorders*, 277, 65–74. doi: 10.1016/j.jad.2020.08.005
- Sackeim, H. A., Brannan, S. K., Rush, A. J., George, M. S., Marangell, L. B., & Allen, J. (2007). Durability of antidepressant response to vagus nerve stimulation (VNS). *International Journal of Neuropsychopharmacology* 10, 817–826. doi: 10.1017/S1461145706007425
- Sackeim, H. A., Prudic, J., Devanand, D. P., Decina, P., Kerr, B., & Malitz, S. (1990). The impact of medication resistance and continuation pharmacotherapy on relapse following response to electroconvulsive therapy in major depression. *Journal of Clinical Psychopharmacology*, 10, 96–104.
- Sayer, N. A., Sackeim, H. A., Moeller, J. R., Prudic, J., Devanand, D. P., Coleman, E. A., & Kiersky, J. E. (1993). The relations between observer-rated and self-report of depressive symptomatology. *Psychological Assessment*, 5, 350–360. doi: 10.1037/1040-3590.5.3.350
- Schwartz, C. E., & Patrick, D. L. (2014). Composite scores in comparative effectiveness research: Counterbalancing parsimony and dimensionality in patient-reported outcomes. *Journal of Comparative Effectiveness Research*, 3, 423–433. doi: 10.2217/ce.14.24
- Sforzini, L. (2021). Lost in translation. The quest for definitions of treatment-resistant depression with a focus on inflammation-related gene expression. *Brain Behav Immun Health*, 16, 100331.
- Shea, M. T., Leon, A. C., Mueller, T. I., Solomon, D. A., Warshaw, M. G., & Keller, M. B. (1996). Does major depression result in lasting personality change? *American Journal of Psychiatry*, 153, 1404–1410. doi: 10.1176/ajp.153.11.1404
- Sherwani, S. I., Khan, H. A., Ekhzaimy, A., Masood, A., & Sakharkar, M. K. (2016). Significance of HbA1c test in diagnosis and prognosis of diabetic patients. *Biomarker Insights*, 11, 95–104. doi: 10.4137/BMI.S38440
- Shih, M. C., Turakhia, M., & Lai, T. L. (2015). Innovative designs of point-of-care comparative effectiveness trials. *Contemporary Clinical Trials*, 45, 61–68. doi: 10.1016/j.cct.2015.06.014
- Sonino, N., Fava, G. A., Raffi, A. R., Boscaro, M., & Fallo, F. (1998). Clinical correlates of major depression in Cushing's disease. *Psychopathology* 31, 302–306. doi: 10.1159/000029054
- Sussman, M., O'Sullivan A. K., Shah, A., Olfson, M., & Menzin, J. (2019). Economic burden of treatment-resistant depression on the U.S. Health Care System. *Journal of Managed Care & Specialty Pharmacy*, 25, 823–835. doi: 10.18553/jmcp.2019.25.7.823
- Thase, M. E., & Howland, R. (1994). Refractory depression: Relevance of psychosocial factors and therapies. *Psychiatric Annals*, 24, 232–240. doi: 10.3928/0048-5713-19940501-09
- Thase, M. E., & Rush, A. J. (1995). Treatment-resistant depression. In F. Bloom & D. Kupfer (Eds.), *Psychopharmacology: The fourth generation of progress* (pp. 1081–1098). New York: Raven.
- Thorpe, K. E., Zwarenstein, M., Oxman, A. D., Treweek, S., Furberg, C. D., Altman, D. G., ... Chalkidou, K. (2009). A pragmatic-explanatory continuum indicator summary (PRECIS): A tool to help trial designers. *Journal of Clinical Epidemiology*, 62, 464–475. doi: 10.1016/j.jclinepi.2008.12.011
- Trivedi, M. H., Fava, M., Wisniewski, S. R., Thase, M. E., Quitkin, F., Warden, D., Ritz, L., ... Team, S. D. S. (2006). Medication augmentation after the failure of SSRIs for depression. *New England Journal of Medicine*, 354, 1243–1252. doi: 10.1056/NEJMoa052964
- Trivedi, M. H., McGrath, P. J., Fava, M., Parsey, R. V., Kurian, B. T., Phillips, M. L., ... Weissman, M. M. (2016). Establishing moderators and biosignatures of antidepressant response in clinical care (EMBARC): Rationale and design. *Journal of Psychiatric Research*, 78, 11–23. doi: 10.1016/j.jpsychires.2016.03.001
- Turkoz, I., Alphas, L., Singh, J., Jamieson, C., Daly, E., Shawi, M., ... Rush, A. J. (2021). Clinically meaningful changes on depressive symptom measures

- and patient-reported outcomes in patients with treatment-resistant depression. *Acta Psychiatrica Scandinavica*, 143, 253–263. doi: 10.1111/acps.13260
- Uher, R., Huzo-Diaz, P., Perroud, N., Smith, R., Rietschel, M., Mors, O., ... Craig, I. (2009). Genetic predictors of response to antidepressants in the GENDEP project. *The Pharmacogenomics Journal*, 9, 225–233. doi: 10.1038/tpj.2009.12
- Uher, R., Perlis, R. H., Placentino, A., Dernovšek, M. Z., Henigsberg, N., Mors, O., ... Farmer, A. (2012). Self-report and clinician-rated measures of depression severity: Can one replace the other? *Depression and Anxiety*, 29, 1043–1049. doi: 10.1002/da.21993
- Verhoeven, J. E., Verduijn, J., van Oppen, P., van Schaik, A., Vinkers, C. H., & Penninx, B. (2020). Getting under the skin: Does biology help predict chronicity of depression? *Journal of Affective Disorders*, 274, 1013–1021. doi: 10.1016/j.jad.2020.05.098
- Vittengl, J. R., Clark, L. A., Thase, M. E., & Jarrett, R. B. (2019). Estimating outcome probabilities from early symptom changes in cognitive therapy for recurrent depression. *Journal of Consulting and Clinical Psychology*, 87, 510–520. doi: 10.1037/ccp0000409
- Wang, P. S., Lane, M., Olfson, M., Pincus, H. A., Wells, K. B., & Kessler, R. C. (2005). Twelve-month use of mental health services in the United States: Results from the national comorbidity survey replication. *Archives of General Psychiatry*, 62, 629–640. doi: 10.1001/archpsyc.62.6.629
- Young, M. (2018). Treatment-resistant depression: The importance of identifying and treating co-occurring personality disorders. *The Psychiatric Clinics of North America*, 41, 249–261. doi: 10.1016/j.psc.2018.01.003
- Zarate, C. A., Jr., & Niciu, M. J. (2015). Ketamine for depression: Evidence, challenges and promise. *World Psychiatry*, 14, 348–350. doi: 10.1002/wps.20269
- Zeier, Z., Carpenter, L. L., Kalin, N. H., Rodriguez, C. I., McDonald, W. M., Widge, A. S., & Nemeroff, C. B. (2018). Clinical implementation of pharmacogenetic decision support tools for antidepressant drug prescribing. *American Journal of Psychiatry*, 175, 873–886. doi: 10.1176/appi.ajp.2018.17111282