Received: 1 March 2019 | Revised: 16 August 2019 | Accepted: 4 October 2019

DOI: 10.1002/da.22968

ANXIETY AND DEPRESSION ASSOCIATION OF AMERICA

WILEY

REVIEW

Defining treatment-resistant depression

Bradley N. Gaynes[1] | Linda Lux[2] | Gerald Gartlehner[2,3] | Gary Asher[4] | Valerie Forman-Hoffman[2] | Josh Green[2] | Erin Boland[2] | Rachel P. Weber[5] | Charli Randolph[5] | Carla Bann[2] | Emmanuel Coker-Schwimmer[5] | Meera Viswanathan[2] | Kathleen N. Lohr[2]

[1]Department of Psychiatry, University of North Carolina at Chapel Hill, Chapel Hill, North Carolina

[2]RTI International, The RTI International-University of North Carolina Evidence-based Practice Center, Durham, North Carolina

[3]Department for Evidence-Based Medicine and Clinical Epidemiology, Danube University, Krems, Austria

[4]Department of Family Medicine, University of North Carolina at Chapel Hill, Chapel Hill, North Carolina

[5]Sheps Center for Health Services Research, The RTI International-University of North Carolina Evidence-based Practice Center, University of North Carolina at Chapel Hill, Chapel Hill, North Carolina

Correspondence
Bradley N. Gaynes, Department of Psychiatry, University of North Carolina at Chapel Hill, 101 Manning Drive, CB #7160, Chapel Hill, NC, 27599.
Email: bradley_gaynes@med.unc.edu

Funding information
Agency for Healthcare Research and Quality, Grant/Award Number: HHSA290201500011I-TO6

Abstract

Background: Varying conceptualizations of treatment-resistant depression (TRD) have made translating research findings or systematic reviews into clinical practice guidelines challenging and inconsistent.

Methods: We conducted a review for the Centers for Medicare & Medicaid Services and the Agency for Healthcare Research and Quality to clarify how experts and investigators have defined TRD and to review systematically how well this definition comports with TRD definitions in clinical trials through July 5, 2019.

Results: We found that no consensus definition existed for TRD. The most common TRD definition for major depressive disorder required a minimum of two prior treatment failures and confirmation of prior adequate dose and duration. The most common TRD definition for bipolar disorder required one prior treatment failure. No clear consensus emerged on defining adequacy of either dose or duration. Our systematic review found that only 17% of intervention studies enrolled samples meeting the most frequently specified criteria for TRD. Depressive outcomes and clinical global impressions were commonly measured; functional impairment and quality-of-life tools were rarely used.

Conclusions: Two key steps are critical to advancing TRD research: (a) Developing a consensus definition of TRD that addresses how best to specify the number of prior treatment failures and the adequacy of dose and duration; and (b) identifying a core package of outcome measures that can be applied in a standardized manner. Our recommendations about stronger approaches to designing and conducting TRD research will foster better evidence to translate into clearer guidelines for treating patients with this serious condition.

KEYWORDS

consensus, definition, guideline, intervention, systematic review, treatment-resistant depression

1 | INTRODUCTION

Patients with major depressive disorder (MDD) or bipolar disorder experience depressive episodes. In 2016 (National Institute of Mental Health, 2017), 6.7% of adults in the United States (16.2 million people) experienced a depressive episode in the past year (Center for Behavioral Health Statistics & Quality, 2016); the great majority are in MDD (Kessler et al., 2003). Treatment for MDD can be inadequate because either the patients do not seek it or the care they receive is substandard

(Wang et al., 2005). Even for MDD patients receiving adequate treatment, only 30% (3% of patients with MDD) experience full recovery or remission. The remaining 70% will either respond without remission (about 20%) or not respond at all (50%; M. H. Trivedi et al., 2006).

Patients whose depressive disorder does not respond satisfactorily to adequate treatment clearly have harder-to-treat depression (Thase & Rush, 1997), generally referred to as treatment-resistant depression (TRD). TRD is a complex phenomenon reflecting a variety of depressive subtypes, psychiatric comorbidity, and coexisting medical illnesses (Berlim, Fleck, & Turecki, 2008). It poses a common, challenging presentation to psychiatric and primary care clinicians (Rush et al., 2006).

TRD represents the highest direct and indirect medical costs among MDD patients (Gibson et al., 2010). Individuals with TRD are twice as likely to be hospitalized; the cost of this hospitalization is more than six times the mean total cost for depressed patients who are not treatment-resistant (Crown et al., 2002). TRD can nearly double both direct and indirect 2-year employer medical expenditures relative to expenditures for patients whose MDD responds to treatment (Ivanova et al., 2010).

Although TRD episodes are most commonly associated with MDD, they are also seen in the depressed phase of bipolar disorder. More than 30% of those suffering from bipolar disorder and receiving treatment do not experience sustained remission of depressive symptoms (Perlis et al., 2006).

No universally accepted operational definition of TRD exists (Trevino, McClintock, McDonald Fischer, Vora, & Husain, 2014). Definitional dilemmas limit the ability of guideline developers or other experts to synthesize information and generalize TRD findings to the array of patient populations encountered in daily practice. Moreover, varying conceptualizations of TRD have made translating research findings or systematic reviews into clinical practice guidelines challenging and inconsistent. Guideline definitions of TRD differ, agreement on what constitutes prior treatment adequacy is lacking, and recommended "next step" interventions can diverge (American Psychiatric Association, 2010; Department of Veterans Affairs and Department of Defense, 2016; Institute for Clinical Systems Improvement, 2016; Lam et al., 2009; National Institute for Health & Clinical Excellence [NICE], 2018).

This review was conducted for the Centers for Medicare & Medicaid Services (CMS) and the Agency for Healthcare Research and Quality (AHRQ). We reviewed definitional and other aspects of TRD in clinical research with two aims: to inform discussions and decisions about how to define the condition and specify the important outcomes measured in research studies, and to clarify how researchers might best design and conduct trials or observational studies to guide clinical practice and health policy. Our original findings, in a detailed report to CMS (Gaynes et al., 2018), addressed literature published through August 17, 2017. This article updates these findings to reflect literature published through July 5, 2019.

2 | MATERIALS AND METHODS

To provide a comprehensive understanding of how various experts and investigators have defined and studied TRD, we employed a mixed-methods approach. For general, contextual questions, our qualitative narrative review gives a broad picture of different definitions, methods, outcome measures, and study designs that researchers, guideline developers, and stakeholders used to address TRD. For questions about specific elements of study design, we conducted a systematic review. We followed standard procedures for systematic literature searches specified in the *AHRQ Methods Guide for Effectiveness and Comparative Effectiveness Reviews* (AHRQ, 2014). The CMS report presents the details of our methods (Gaynes et al., 2018).

In this article, we examine five key questions (KQs).

Narrative review: KQs 1–3

1. What definitions of TRD appear in these sources? Do definitions converge on the best one?
2. What methods do investigators use to diagnose this condition in clinical research? Does a consensus exist about the best ways to reach a clear diagnosis?
3. What measures (i.e., endpoints or outcomes) exist to determine the success or failure of treatment in TRD studies? What clinical focus do they represent (e.g., severity)? What psychometric and other properties do they have?

Systematic review: KQs 4 and 5

1. What are the inclusion criteria for patients in TRD studies, specifically patient characteristics, prior treatments, and diagnostic characteristics?
2. How do these criteria compare or contrast with definitions from the narrative review?

2.1 | Inclusion and exclusion criteria

Our population of interest was adults 18 years of age or older with depression who had not responded to treatment(s). The depressive illness could be part of either MDD or bipolar disorder; these diagnoses had to be per the *Diagnostic and Statistical Manual of Mental Disorders (DSM)*, fourth edition (DSM, 1994; American Psychiatric Association, 1994) or fifth edition (DSM, 2013; American Psychiatric Association, 2013). Eligible interventions included those that had been tested as a treatment targeting TRD in adults *and* identified by guidelines, consensus statements, the Medicare Evidence Development & Coverage Advisory Committee (MEDCAC) panel of April 27, 2016, or systematic reviews as alternatives for TRD treatment studies. These criteria ensured consideration of interventions with a minimum threshold amount of data addressing their effectiveness in TRD populations. Comparison groups included concurrent control groups (e.g.,

active, sham, or placebo) and a control group from an interrupted time series. Table 1 provides an abbreviated list of PICOTS (populations, interventions, comparators, outcomes, time frames, and settings); online supplement documents detailed inclusion/exclusion criteria and a PRISMA flow diagram.

2.2 | Literature search

We systematically searched the published literature from January 1, 1995, through July 5, 2019, that was indexed in MEDLINE®, EMBASE, PsycINFO, and Cochrane Library *and* that addressed the treatment of TRD in adults. The aim was to assemble literature relevant to current definitions of TRD with diagnoses consistent with definitions in either of the DMS editions noted above. For the two systematic KQs, we set the earliest publication date for intervention studies as January 1, 2005.

We also searched for consensus statements, management guidelines, and relevant government materials from various federal agencies: CMS (and MEDCAC), Food and Drug Administration (FDA), National Institutes of Health, National Institute of Mental Health, and the Substance Abuse and Mental Health Services Administration. We abstracted information relevant to KQs 1 through 3 from such sources. We also searched other websites such as Clinicaltrials.gov, Guideline.gov (AHRQ's National Guidelines Clearinghouse), HSRProj (Health Services Research Projects in Progress database), and UpToDate.

TABLE 1 Inclusion criteria (abbreviated) for literature searches

PICOTS	Inclusion criteria
Population	All adult populations (≥18-years old) identified as having a primary diagnosis of depression (including MDD and bipolar disorder) who have had a depressive episode and have not responded to treatment(s) (the least stringent definition of TRD). The depressive episode must be part of an MDD or a bipolar disorder.
Interventions	Any pharmacologic intervention tested as a treatment for TRD as primary therapy or as an augmentation agent to an existing primary therapy.[a] Any nonpharmacologic device or procedure tested as a treatment for TRD as primary therapy or as augmentation to an existing primary therapy.[b] Psychotherapy (i.e., cognitive behavioral therapy, third-wave cognitive behavioral therapy, psychodynamic therapies, and integrative therapies). CAM interventions and formal exercise programs.
Comparators	All those above in studies with concurrent control groups or control groups from an interrupted time series or pre/post studies with interrupted time series.
Outcomes	Benefits that are reported as primary endpoints (or outcomes). Reduction in suicidal ideation or suicide attempts. Quality of life. Response to treatment. Remission. Change in depressive severity. Functional capacity (physical and cognitive functioning measured by validated scales). Speed of remission. Speed of response. Intervention durability (rates or counts of recurrence of a depressive episode for those who have remitted).
Timing	Any study duration.
Setting	Studies in very highly developed countries (United Nations Development Programme, 2015).[c]
Study designs	For KQs 1–3: Consensus statements, guidelines, or other materials, and systematic reviews. For KQs 4 and 5: Randomized or prospective nonrandomized, or observational studies (including concurrent controls and interrupted time series).

Abbreviations: CAM, complementary and alternative medicine; KQ, key question; MDD, major depressive disorder; TRD, treatment-resistant depression.
[a]Pharmacologic: SSRIs: citalopram, escitalopram, fluoxetine, fluvoxamine, paroxetine, and sertraline; SNRIs: desvenlafaxine, duloxetine, levomilnacipran, mirtazapine, and venlafaxine; NADRIs: bupropion; TCAs: amitriptyline, clomipramine, desipramine, doxepin, imipramine, maprotiline, and nortriptyline; MAOIs: phenelzine, selegiline transdermal, and tranylcypromine; 5HT Ras: nefazodone, trazodone, vilazodone, and vortioxetine; Atypical antipsychotics: cariprazine and quetiapine; NMDAs: ketamine; Other pharmacologic for combination or augmentation: Atypical antipsychotics: aripiprazole, asenapine maleate, brexpiprazole, clozapine, iloperidone, lurasidone, olanzapine, paliperidone, quetiapine, risperidone, and ziprasidone; Anticonvulsants: carbamazepine, lamotrigine, oxcarbazepine, topiramate, and valproic acid; Psychostimulants: amphetamine-dextroamphetamine, armodafinil, dexmethylphenidate, dextroamphetamine, lisdexamfetamine, methyphenidate, and modafinil; Mood stabilizers: lithium and divalproex; Other augmenters: bupropion, buspirone, clonidine, liothyronine, pindolol, pramipexole, and triiodothyronine (T3).
[b]BST, deep-brain stimulation, electroconvulsive therapy; CES, cranial electrotherapy stimulation; ECT, electroconvulsive therapy; MST, magnetic seizure therapy; rTMS, repetitive transcranial magnetic stimulation; tDCS, transcranial direct stimulation; VNS, vagus nerve stimulation.
[c]"Very High" on Human Development Index: Andorra, Argentina, Australia, Austria, Bahrain, Belgium, Brunei Darussalam, Canada, Chile, Croatia, Cyprus, Czech Republic, Denmark, Estonia, Finland, France, Germany, Greece, Hong Kong China (SAR), Hungary, Iceland, Ireland, Israel, Italy, Japan, Korea (Republic of), Kuwait, Latvia, Liechtenstein, Lithuania, Luxembourg, Malta, Montenegro, Netherlands, New Zealand, Norway, Poland, Portugal, Qatar, Saudi Arabia, Singapore, Slovakia, Slovenia, Spain, Sweden, Switzerland, United Arab Emirates, United Kingdom, United States (United Nations Development Programme, 2015).

Further, the Scientific Resource Center for AHRQ's Evidence-based Practice Center Program contacted relevant stakeholders, including manufacturers of prescription medications and medical devices used to treat MDD, for scientific information packets with any unpublished information on the efficacy and/or safety of their products when used specifically to treat TRD patients.

2.3 | Literature review

Trained members of the research team dually reviewed all titles and abstracts for eligibility based on the pre-established criteria (Table 1). Studies marked for possible inclusion by either reviewer underwent full-text review, as did any study with inadequate information in its abstract. These researchers dually reviewed each full-text article for inclusion or exclusion against the eligibility criteria. We documented reasons for exclusion at this stage; we also tagged studies selected for inclusion with the relevant KQ(s) that the article addressed. Disagreements about inclusion were resolved by consensus discussion.

2.4 | Data abstraction

We abstracted data relevant to the KQs from any studies meeting inclusion criteria into a standardized template. One member of the research team collected the data; another (senior) investigator reviewed the abstraction for accuracy and completeness.

2.5 | Data synthesis

All synthesis of the literature was done narratively.

3 | RESULTS

Our searches identified 3,269 citations, of which we included 260 articles: 41 for KQs 1 through 3, and 219 for KQs 4 and 5. Results from a single trial might be reported in several articles. Included articles for KQs 4 and 5 summarized 185 unique studies: 163 randomized controlled trials (RCTs), 4 nonrandomized trials, 15 observational studies, 2 case-control studies, and 1 interrupted time series.

3.1 | Narrative review: the current understanding of treatment-resistant depression

3.1.1 | Definitions of treatment-resistant depression (KQ 1)

We identified four categories of TRD definitions, distinguished primarily by the number of previous failed antidepressant treatment attempts (Table 2). Identified this way, individuals either did or did not have TRD. The great majority of definitions of TRD address MDD, for which there were three: one or more, two or more, and three or more treatment failures. For bipolar patients, we found a systematic review and a guideline that used one or more failures for TRD and one recent consensus statement that recommends two or more.

Also, as shown in Table 3, we identified five TRD staging models (which delineate the degree of resistance along a spectrum); research about reliability and validity was limited. The Antidepressant Treatment History Form (ATHF; Oquendo et al., 2003) is a 1-to-5 scale; it ranks medication resistance according to the adequacy of prior antidepressant treatment (indicated by medication dose, treatment duration, and patient adherence). Prediction of treatment response is limited to studies with electroconvulsive therapy (ECT); three prospective studies showed an association between a high score on the ATHF and worse patient outcomes. Reliability was good in two studies (Ruhé, van Rooijen, Spijker, Peeters, & Schene, 2012). This tool appears to be intended primarily for use in research settings (Washington State Health Care Authority, 2014).

The Thase and Rush Staging Model (TRSM; Thase & Rush, 1997) proposes a five-part categorical scale; it stages patients according to the number of classes of antidepressants that have failed to provide a response (higher stages are more treatment-resistant). Treatment resistance moves from more frequently used antidepressants (such as selective serotonin reuptake inhibitors or tricyclic antidepressants) to less frequently used therapies (e.g., monoamine oxidase inhibitors or ECT). The predictive value of the TRSM has not been systematically assessed; reliability has not been tested (Ruhé et al., 2012).

Two models are variations on TRSM. The European Staging Model (Souery et al., 1999) distinguishes among three groups: nonresponse, TRD, and chronic resistant depression (Souery et al., 1999). "Nonresponders" are patients who do not respond to one form of treatment. Patients are considered treatment-resistant after they have a poor response to a second treatment attempt with a different class of antidepressants. Further staging depends on the duration of treatment with adequate medication doses. TRD staging ranges from 1 to 5; resistance beyond 12 months indicates a distinct category—chronic resistant depression. No studies tested the predictive utility or reliability of this model (Ruhé et al., 2012).

Another model related to the TRSM is the Massachusetts General Hospital Staging model (MGH-s). It is a function primarily of the number of prior antidepressant failures (with no hierarchy of antidepressant classes; Fava, 2003). It also considers the optimization of treatments, augmentation and combination strategies, and prior failed ECT. Patients receive certain points for these components. The MGH-s produces a continuous score, reflecting the level of treatment resistance. A retrospective chart review showed an association between a higher MGH-s score and worse outcome (Ruhé et al., 2012). No study has assessed reliability.

The fifth staging model, the Maudsley Staging Method (MSM; Fekadu et al., 2009), summarizes the TRD stage into a single score, ranging from 3 to 15. Like the other models, MSM considers the number of treatment failures, and it considers augmentation strategies and ECT treatment (as does the MGH-s). Unlike the TRSM, European Staging Model, and MGH-s, however, the MSM

TABLE 2 Four categories of definitions of treatment-resistant depression by the number of treatment failures and components of definition

Number of treatment failures	Type of publication on TRD treatments, date	Defines nonresponse or lack of remission	Specifies current episode	Defines adequate dose	Defines adequate duration (weeks)
For MDD TRD: One or more	Seminal article on defining TRD, 1996 (Fava & Davidson, 1996)	+	−	+	≥6
	SR—Pharmacologic, 2007 (Berlim & Turecki, 2007)	+/−−	+/−−	+/−−	≥4 to ≥8
	SR—Lamotrigine augmentation, 2010 (Thomas, Nandhra, & Jayaraman, 2010)	−	−	−	4
	SR—Psychotherapy, 2011 (R. B. Trivedi, Nieuwsma, & Williams, 2011)	−	−	+/−−	≥6
	Review defining TRD, 2014 (Trevino et al., 2014)	−	−	+	6 to 8
	SR—rTMS, 2015 (Silverstein et al., 2015)	−	−	−	−−
	SR—rTMS, 2015 (Zhang et al., 2015)	−	+	−	4 to 6
	SR—Predictors of nonresponse, 2016 (de Carlo, Calati, & Serretti, 2016)	−	−	−	−
	SR—Brexpiprazole augmentation, 2017 (Yoon et al., 2017)	−	+	−	6 to 12
	SR—Psychotherapy, 2018 (Ijaz et al., 2018)	−	−	+	≥4
Two or more	Seminal article on definition of TRD, 2001 (Sackeim, 2001)	+	+	+	≥4
	SR—Pharmacologic treatments, 2007 (Berlim & Turecki, 2007)	++/−	+/−−	++/−	4 to 6
	ICER coverage policy analysis, 2012 (New England Comparative Effectiveness Public Advisory Council, 2012)	+	+	−	−
	SR - lithium or atypical antipsychotics, 2013 (Edwards, Hamilton, Nherera, & Trevor, 2013)	−	+/−−	−	≥4
	SR—rTMS, 2014 (Gaynes et al., 2014)	+	+/−−	+	≥4
	SR—nonpharmacological, 2014 (Washington State Health Care Authority, 2014)	+/−−	+/−−	+	≥4
	Australian/New Zealand Clinical Practice Guideline, 2015 (Malhi et al., 2015)	−	−	+	≥4 to 8
	VA/DoD Clinical Practice Guideline, 2016 (Department of Veterans Affairs & Department of Defense, 2016)	−	−	+	≥4 to 6
	SR—Pharmacologic and somatic, 2017 (Papadimitropoulou, Vossen, Karabis, Donatti, & Kubitz, 2016)	−	+/−−	−	−
	SR—Psychological and pharmacologic (Strawbridge et al., 2019)	+	+	−	++/−
	SR—Asia Pacific perspectives on definition of TRD (Ng et al., 2019)	+	−	+	6 to 8
Three or more	ICSI Adult Depression in Primary Care Guideline, 2016 (Institute for Clinical Systems Improvement, 2016)	+	−	−	−
For bipolar TRD: One or more	SR—Nonpharmacological, 2014 (Washington State Health Care Authority, 2014)	+	−	−	10 to 12
	Australian/New Zealand Clinical Practice Guideline, 2015 (Malhi et al., 2015)	−	−	+	≥3
	Consensus definition of UK bipolar experts (Hidalgo-Mazzei et al., 2019)	+	+	+	8

Abbreviations: ICER, Institute for Clinical and Economic Review; ICSI, Institute for Clinical Systems Improvement; MDD, major depressive disorder; rTMS, repetitive transcranial magnetic stimulation; SR, systematic review; TRD, treatment-resistant depression; VA/DoD, Department of Veterans Affairs/Department of Defense; +, definition was provided; −, definition was not provided; ++/−, more studies in review provided definition; +/−−, more studies in the review did not provide definition.

TABLE 3 Staging models for treatment-resistant depression to define the spectrum of illness

Models	How severity scored	Failure defined	Specify current episode	Define adequate dose	Define adequate duration	Staging schema	Predictive validity and reliability tested? Other Comments
Antidepressant Treatment History Form (Ruhé, van Rooijen, Spijker, Spijker, & Schene, 2012)	Sum score based on points per treatment	−	+	+	≥4 weeks	Five stages 0–5)	Predictive validity confirmed in three prospective ECT studies; reliability good in two studies. Does not correspond with the number of treatment failures; does not count psychotherapy in failed trials
Thase and Rush Staging Model (TRSM; Berlim & Turecki, 2007; Fava, 2003; Ruhé et al., 2012; Thase & Rush, 1997)	Stages, with higher-numbered stages indicating a greater degree of treatment resistance	+	−	−	≥4 weeks	Five stages 1–5)	Predictive value has not been systematically assessed; reliability has not been tested. Stage II corresponds with two treatment failures; considers the number of classes of ADs that have failed to provide a response but not psychotherapy
European Staging Model (Fekadu et al., 2009; Ruhé et al., 2012; Souery et al. 1999)	Number of weeks with treatment resistance	+	+	−	Varies by nonresponder, TRD, and CRD	Three categories: nonresponder, TRD, and CRD	Predictive value has not been systematically assessed; reliability has not been tested All TRD stages are consistent with two treatment failures
Massachusetts General Hospital Staging model (MGH-s; Fava, 2003; Ruhé et al., 2012)	Points based on the number of prior failures	+	−	+	≥6 weeks	Three stages based on the number of AD failures and three points for an ECT failure	Retrospective chart review showed an association between higher MGH-s score and worse outcome; a retrospective study showed the MGH-s model better predicted nonremission than the TRSM. Reliability for these models was not tested. No direct correspondence with two or more treatment failures.
Maudsley Staging Model (Fekadu et al., 2009; Ruhé et al., 2012)	Points per number of prior attempts, duration, symptoms severity, augmentation use, ECT	+	+	+	Varies by intervention	Points based on duration, symptom severity, number of treatment failures, augmentation, ECT	The only tool with prospective testing showing good validity; two studies showed the MSM score predicted future nonresponse significantly better than the TRSM. Reliability not tested. No direct correspondence with two or more treatment failures

Abbreviations: AD, antidepressant; CRD, chronic resistant depression; ECT, electroconvulsive therapy; RCT, randomized controlled trial; rTMS, repetitive transcranial magnetic stimulation; TRD, treatment-resistant depression; +, definition was provided; −, definition was not provided.

includes two disease characteristics, namely, duration and symptom severity at baseline. This model is the only one with validity assessed with prospective data (Fekadu et al., 2009). Higher scores were associated with failure to achieve remission; the model correctly predicted treatment resistance in more than 85% of cases. Reliability testing has not been reported.

These models appeared to be equally valid for documenting treatment failure in depressed patients. Their applicability and feasibility in clinical practice, however, are unclear (Table 3). Identified this way, people could have TRD along a spectrum of severity.

No consensus exists on the best TRD definition. However, the majority of systematic reviews and guidelines or consensus statements reported that the commonly used definitions were based on treatment of patients whose depression failed to respond (a decrease in depressive severity of at least half) or did not go into remission (complete recovery as measured by a score on a depressive severity instrument below a threshold) following two or more treatment attempts of an adequate dose and duration.

Experts do not agree on how to define an adequate dose and adequate duration. Typically, the minimum duration cited is 4 weeks.

3.1.2 | Preferred diagnostic tools (KQ2)

The literature emphasizes that a diagnosis of a major depressive episode as a part of MDD or bipolar disorder can be made through a standard clinical evaluation (based on the 1994 *DSM-IV* or 2013 *DSM-5*, International Classification of Diseases, or Research Diagnostic Criteria) or through a more structured clinical assessment (Mini International Neuropsychiatric Interview, Structured Clinical Interview for the Diagnostic and Statistical Manual, or Schedule for Affective Disorders and Schizophrenia). Nevertheless, no consensus exists about the preferred approach for diagnosing TRD.

Diagnosing TRD entails one of two further steps: (a) collecting a careful history before treatments (e.g., the number of prior pharmacologic attempts of adequate dose and duration that did not produce remission) or (b) administering a structured, staging tool (e.g., the ATHF, TRSM, European Staging Model, MGH-s, or the MSM) to confirm treatment resistance. No preferred approach exists, the evidence base is limited, and careful history has not been compared directly with a structured tool. Studies do not show any relationship between the medical setting and choice of a diagnostic tool; the selection may rest on feasibility.

3.1.3 | Preferred outcome measures to determine success or failure (KQ3)

No consensus exists about the best outcome measure to use for TRD. The three main categories of outcome measures—depression-specific measures, general psychiatric status measures, and functional scales have both patient-reported and clinician-administered versions available. The Hamilton Depression Rating Scale (HAM-D) is the most common depression-specific measure. Remission (complete recovery as measured by a score below a threshold) is the preferred endpoint regardless of the tool.

General psychiatric status measures were infrequently described; most commonly reported was the Clinical Global Impression scale. Various functional scales have been reported; no one scale is the most frequently used. Most measures have adequate psychometric properties (e.g., reliability and validity) for measuring depressive outcomes. The minimum clinically important difference has been variably defined for many of these outcome measures; none is a consensus preference.

3.2 | Systematic review: how controlled trials investigate treatment-resistant depression

We divided interventions tested in the 185 unique studies into four categories: (a) brain stimulation treatments (BST), which included ECT, repetitive transcranial magnetic stimulation (rTMS), vagal nerve stimulation, and deep brain stimulation (85 studies); (b) pharmacotherapy, including ketamine (77 studies); (c) psychotherapy (15 studies); and (d) complementary and alternative medicine (CAM) therapies and exercise (8 studies).

3.2.1 | Inclusion criteria for intervention studies (KQ 4)

Confirmation of TRD for study entry was often poorly described. The HAM-D and the Montgomery–Åsberg Depression Rating Scale were commonly used to set minimum depressive severity thresholds for study entry; most studies involved patients with moderately severe depression. Studies were inconsistent about the necessary duration of prior treatment attempts for study entry. Most (155/185) studies required at least one, and often two, prior failed treatment attempts of adequate therapy. Several patient characteristics were considered (albeit only rarely) for study entry: duration of depressive symptoms (n = 30), prior augmentation or combination therapy (n = 13), prior treatment intolerance (n = 9), prior psychotherapy (n = 7), and suicidality (n = 3).

3.2.2 | Inclusion criteria compared with definitions of treatment-resistant depression (KQ 5)

Inclusion criteria as specified by the 185 eligible TRD studies did not align closely with the definition(s) of TRD identified in the narrative review. The most common TRD definition from our narrative review involved a minimum of two failed prior adequate antidepressant studies. By contrast, the most common definition from included intervention studies was a minimum of one failed trial used in 89 studies (48%); only 66 (37%) required a minimum of two failed trials. (Table 4).

We assessed how well inclusion criteria matched what experts identified as adequate dose and duration. For adequate dose, we first identified whether studies stated that they had restricted eligibility to patients who had received an adequate prior dose as

TABLE 4 Numbers of studies of treatment-resistant depression considering or confirming key inclusion criteria for defining the diagnosis

Minimum prior treatment failures	Number of studies using this minimum for inclusion	Adequate dosage considered	Adequate dosage confirmed	Adequate duration considered	Adequate duration confirmed	TRD staged
1	89	78	66	78	72	23
2	68	52	33	55	46	27
3	9	8	6	8	5	5
≥4	10	5	4	6	3	3
Not considered	9	4	7	4	4	9
Total	185	146	112	150	128	59

Abbreviation: TRD, treatment-resistant depression.

part of their inclusion criteria (i.e., that an adequate dose was *considered* in determining eligibility). If so, we subsequently identified whether the study had systematically *confirmed* this dose by specifying dosage levels through interview, questionnaire (e.g., Antidepressant Treatment Response Questionnaire), or other formal clarification. We considered a statement that eligibility criteria required a minimum therapeutic dose as stated by product labeling to indicate confirmation.

Similarly, for the adequate duration, we first identified whether the studies stated that they had *considered* this criterion by restricting eligibility to those patients with a prior adequate duration. If so (for such studies), we then determined whether the investigators had *confirmed* the duration: that is, clarified that patients previously received what KQ 1 had indicated was an adequate dose. In KQ 1, approximately one-half of the eligible reviews and guidelines identified a minimum of 4 weeks of treatment; the other half identified it as 6 weeks. We defined an adequate dose here as 4 weeks because one primary tool to confirm the adequacy of dose and duration, the ATHF, required at least 4 weeks to be considered as adequate duration.

Of 185 studies, 146 (79%) *considered* in their selection criteria whether the patient had been treated previously with an adequate dose; 112 (61%) systematically *confirmed* that the dose was adequate by specifying dosage levels. Of all 185 studies, 150 (81%) *considered* in their selection criteria whether prior treatments were of adequate duration; 128 (69%) systematically *confirmed* that the duration was adequate (≥4 weeks of treatment). Fifty-nine studies (32%) set inclusion criteria based on the stage of TRD using a staging model.

Few trials defined the TRD population according to what most conceptualized as the current definition of TRD. Only 19% of studies had *all* three commonly described criteria for TRD: a minimum of two prior treatment failures, confirmation that a dose was adequate, and confirmation that duration was 4 weeks or longer.

4 | DISCUSSION

The concept of TRD has generated a substantial amount of study and collaboration that has attempted to characterize TRD. For example, in Europe, the European Group for the Study of Resistant Depression has published several studies and reviews (Bartova et al., 2019; Kautzky et al., 2019; Schosser et al., 2012; Souery et al., 2007) that have influenced current treatment recommendations of the European Medicines Agency (European Medicines Agency, 2013). We assessed how well this available literature addressed two main questions: what was the consensus definition for TRD, and how this definition applied in intervention trials.

Our narrative review indicated that no consensus definition existed for TRD. We identified four basic definitions for TRD (three for MDD, one for bipolar disorder). The most common TRD definition for MDD requires a minimum of two prior treatment failures and confirmation of prior adequate dose and duration. No clear consensus emerged on how to define adequacy of either dose or duration. We identified little consensus about the best tools to diagnose TRD or measure its outcome.

Our systematic review indicated that inclusion criteria as specified by the eligible TRD trials or observational studies generally did not closely align with TRD definitions from the narrative review. Only 20% of studies defined TRD as two prior treatment failures *and* confirmation of prior adequate dose and duration. Depressive outcomes were the frequently reported endpoints; clinical global impressions were also often assessed. Tools to measure functional impairment and quality of life were infrequently used.

4.1 | Findings in relationship to what is already known

The variability in the definitions and conceptualization of TRD is consistent with other reports from the past decade identifying the lack of any standard, systematic definition of TRD (Berlim & Turecki, 2007; European Medicines Agency, 2013; Ruhé et al., 2012; Trevino et al., 2014). Defining the adequacy of dose and duration, clarifying failure (as remission or response and after what length of time), and determining whether TRD requires prior use of different classes of antidepressants are all variably defined and implemented.

Taken together, the available literature highlights the resulting difficulty in synthesizing information across trials or other types of studies or documents. This characteristic of the evidence base also underscores the problems of translating research findings into guidelines for selecting better treatment options for TRD patients.

Our systematic review highlighted some new key findings. The mismatch between the most common number of treatment failures (at least two) and what the most recent literature has assessed (at least one failure) was stark. The failure of inclusion criteria in recent TRD studies to confirm systematically an adequate dose (61%) or duration (70%) has not previously been described. Finding that only 19% of recent intervention studies are consistent with the most common definition of TRD underscores the inconsistency in TRD definitions. All these results highlight another concern: how to compare and synthesize data across treatment studies.

Finally, despite the substantial morbidity associated with TRD, the relative infrequency of the use of patient-oriented outcomes such as functional impairment and quality-of-life measures in considering the benefits of TRD treatment was newly demonstrated. So too was the infrequent measurement of both adherence to treatment and healthcare services use.

4.2 | Implications for clinical and policy decisionmaking

This current state of evidence underscores the challenges facing clinicians. Effective treatments exist, but because of the variability in TRD definitions and study populations, determining to which patients the research results apply is difficult and could lead to mistakes in treatment. For example, treatments that may work best for those who received two ineffective treatments may not work best for those who received five.

Similarly, the state of the evidence poses difficulties for policymakers. Officials, at public agencies or private-sector organizations, must be confident that two main assumptions are met. The first is that the population of patients with TRD is being consistently and systematically defined; the second is that meaningful and comparable outcomes of importance to both patients and clinicians are being monitored. Neither assumption is consistently met in the literature, limiting translation of this treatment information into actual care.

The high level of morbidity associated with TRD is clear. For adequate clinical and policy decisionmaking about TRD patients, however, a widely agreed-upon definition of the condition that addresses how best to determine the number of prior treatment failures and the adequacy of dose and duration is critical. Some means of systematically monitoring this for TRD on a large scale is warranted; this could be, for instance, a treatment registry using common data elements. Such a resource could substantially help clarify which criteria best define TRD, what the course of illness is, and how interventions might affect that course.

4.3 | Research recommendations

We propose several steps to address existing evidence gaps and substantially improve the study and treatment of patients with TRD.

Reducing the heterogeneity of how TRD patient populations are defined is a necessary first step. Perhaps the most critical task is to reach an agreement on a standardized, systematic, and feasible definition of TRD. Such a definition should clearly specify the number of prior treatment attempts, an adequate dose, and an adequate duration. At the very least, the minimum number of past failed therapy attempts should be two (regardless of the class of antidepressants). Another necessary part of this "definitional" step for researchers is standardized confirmation of the adequacy of prior treatment attempts.

Systematic accounting for potential confounders is also crucial. Six critical factors must be considered: depressive severity, duration of current episode, prior treatment intolerance, prior augmentation or combination therapy, prior psychotherapy, and psychiatric comorbidities. Other sociocultural factors that may influence the course of TRD (e.g., sex and age) also deserve further study, as may genetic factors (Malhi et al., 2015). Randomization can account for some measured and unmeasured confounders in larger trials. However, the smaller RCTs that we identified had imbalances in baseline characteristics and rarely adjusted for such differences; nonrandomized TRD studies adjusted for potential confounders less than half of the time.

Agreement on a package of outcome measures to be administered in a standard way should be strongly encouraged. The field would benefit from an evidence-informed, multistakeholder consensus process to develop a core outcome set for TRD, potentially something similar to the Outcome Measures in Rheumatology (OMERACT; https://www.omeract.org/). Of particular importance is including one measure of the following: depressive severity, general psychiatric status, functional impairment, quality of life, and adherence to medications or other interventions. Another key outcome to measure is suicidality (to assess for a reduction in suicidality).

In addition, common use of measures will allow for better comparisons among trials. These will improve our ability to combine studies for meta-analyses. Patient-reported instruments may be preferred because they are more feasible, generally speaking, and more patient-centered than clinician-reported instruments.

Researchers and clinicians should agree on a standard length of treatment. The key is to provide enough time for patients to receive an adequate dose and duration of the intervention. Given the chronicity of TRD and the time to reach an adequate dose and length of treatment, at least 2 months is the bare minimum for the study duration. Also, the increasing risk of relapse with greater levels of resistance (Rush et al., 2006) argues for demonstrating longer-term efficacy for the more treatment-resistant patients, suggesting a need for even longer trials for the more severely resistant. For this latter group, who may have a more chronic and persistent course, different treatments might be required than for those whose depressive episodes are more episodic.

We found only a very few studies of TRD interventions other than pharmacological or BST interventions (i.e., psychotherapies and CAM or exercise). This gap reduced the evidence base relevant for patients who prefer to avoid, or for whom it would be inappropriate to try, pharmacological agents or more invasive procedures. Considering less-studied interventions could help

inform patient decisions about options and improve the level of shared or informed decisionmaking.

Trials or robust types of observational studies to test the *effectiveness* of all such interventions in real-world settings are necessary. Targeting only efficacy (via RCTs) may produce information for clinicians, patients, or policymakers that cannot easily be applied in "ordinary," every-day circumstances.

To allow for better assessment of quality in TRD, publications of RCTs need to adhere to Consolidated Standards of Reporting Trials specifications for reporting (Boutron et al., 2017). Similarly, publications of nonrandomized controlled trials or observational studies should adhere to Strengthening the Reporting of Observational Studies in Epidemiology (STROBE; von Elm et al., 2008). Documenting all steps in such investigations, reporting on all planned outcomes, and otherwise ensuring complete transparency for this study are critical actions in adding to the professional literature. These steps would help ensure a consistent definition of TRD and its reported outcomes.

Finally, TRD needs to be monitored, consistently, systematically, and on a large scale. For instance, a treatment registry using common data elements could substantially help clarify the criteria that best define TRD, what the course of illness is, and how interventions might affect that course. Coordination among different specific treatment registries that already exist (e.g., the vagal nerve stimulation registry that the FDA requires (Aaronson et al., 2017; Naumann et al., 2017) and the transcranial magnetic stimulation registry recently launched by Neurostar (NeuroStar Advanced Therapy, 2017) and that have been suggested (e.g., a ketamine registry; Singh, Singh, Kar, & Pahuja, 2017) would be a necessary step. Data quality is a key challenge for such an enterprise.

5 | CONCLUSION

We encountered substantial diversity at every stage of research on TRD interventions. Of particular concern was the lack of consensus about various elements of even a TRD diagnosis and appropriate inclusion or exclusion criteria. Little or no agreement about important outcomes and how to assess them hampered analysis. An extensive set of recommendations about additional and more robust approaches to the design and conduct of this study will foster better evidence to translate into clearer guidelines for treating patients with this serious condition.

ACKNOWLEDGMENTS

The authors gratefully acknowledge the following individuals for their contributions to this project and deeply appreciate their considerable support, commitment, and contributions. From the Agency for Healthcare Research and Quality: Aysegul Gozu, M.D., M.P.H., our task order officer, and Elise Berliner, Ph.D., director of the Technology Assessment Program at the Center for Evidence and Practice Improvement. From the Center for Medicare & Medicaid Services' Coverage and Analysis Group: Susan M. Miller, M.D.; Rosemarie Hakim, Ph.D.; Linda Gousis, J.D.; and Lori Ashby, M.A., director of the Division of Medical and Surgical Services. From the RTI International–University of North Carolina Evidence-based Practice Center (EPC) staff: Carol Woodell, B.S.P.H., EPC project manager; Dan Jonas, M.D., M.P.H., EPC co-director; and Lynn Whitener, Dr. P.H., M.S.L.S., librarian. Agency for Healthcare Research and Quality (AHRQ), Rockville, MD (Contract No. HHSA290201500011I_HHSA29032010T). The findings and conclusions in this document are those of the authors, who are responsible for its contents; the findings and conclusions do not necessarily represent the views of AHRQ or CMS. Therefore, no statement in this report should be construed as an official position of AHRQ, CMS, or of the U.S. Department of Health and Human Services.

ORCID

Bradley N. Gaynes http://orcid.org/0000-0002-8283-5030
Linda Lux http://orcid.org/0000-0003-2560-7362

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*(7), 640–648.

Agency for Healthcare Research and Quality (2014). Methods guide for effectiveness and comparative effectiveness reviews (Report No. 10(14)-EHC063-EF). Rockville, MD; AHRQ Publication.

American Psychiatric Association (1994). *Diagnostic and Statistical Manual of Mental Disorders*. Washington, DC: American Psychiatric Association.

American Psychiatric Association (2010). In Workgroup on Major Depressive Disorder (Ed.), *Practice guidelines for the treatment of patients with major depressive disorder*. Washington, DC: American Psychiatric Association.

American Psychiatric Association (2013). *Diagnostic and Statistical Manual of Mental Disorders, DSM-5* (5th ed.). Arlington, VA: American Psychiatric Association.

Bartova, L., Dold, M., Kautzky, A., Fabbri, C., Spies, M., Serretti, A., … Kasper, S. (2019). Results of the European Group for the Study of Resistant Depression (GSRD)—Basis for further research and clinical practice. *The World Journal of Biological Psychiatry*, *20*, 427–448.

Berlim, M. T., Fleck, M. P., & Turecki, G. (2008). Current trends in the assessment and somatic treatment of resistant/refractory major depression: An overview. *Annals of Medicine*, *40*(2), 149–159.

Berlim, M. T., & Turecki, G. (2007). What is the meaning of treatment-resistant/refractory major depression (TRD)? A systematic review of current randomized trials. *European Neuropsychopharmacology*, *17*(11), 696–707.

Boutron, I., Altman, D. G., Moher, D., Schulz, K. F., Ravaud, P., & Group, C. N. (2017). CONSORT statement for randomized trials of nonpharmacologic treatments: A 2017 update and a CONSORT extension for nonpharmacologic trial abstracts. *Annals of Internal Medicine*, *167*(1), 40–47.

de Carlo, V., Calati, R., & Serretti, A. (2016). Socio-demographic and clinical predictors of non-response/non-remission in

treatment-resistant depressed patients: A systematic review. *Psychiatry Research*, *240*, 421–430.

Center for Behavioral Health Statistics and Quality (2016). *Results from the 2015 National Survey on Drug Use and Health: Detailed Tables*. Rockville, MD: Substance Abuse and Mental Health Services Administration.

Crown, W. H., Finkelstein, S., Berndt, E. R., Ling, D., Poret, A. W., Rush, A. J., & Russell, J. M. (2002). The impact of treatment-resistant depression on health care utilization and costs. *Journal of Clinical Psychiatry*, *63*(11), 963–971.

Department of Veterans Affairs, Department of Defense (2016). *VA/DoD Clinical Practice Guideline for the Management of Major Depressive Disorder*. Washington, DC: Department of Veterans Affairs and Department of Defense.

Edwards, S., Hamilton, V., Nherera, L., & Trevor, N. (2013). Lithium or an atypical antipsychotic drug in the management of treatment-resistant depression: A systematic review and economic evaluation. *Health Technology Assessment*, *17*(54), 1–190.

von Elm, E., Altman, D. G., Egger, M., Pocock, S. J., Gøtzsche, P. C., VandenBroucke, J. P., & Initiative, S. (2008). The Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) statement: Guidelines for reporting observational studies. *Journal of Clinical Epidemiology*, *61*(4), 344–349.

European Medicines Agency (2013). *Guideline on clinical investigation of medicinal products in the treatment of depression*. London, UK: European Medicines Agency.

Fava, M. (2003). Diagnosis and definition of treatment-resistant depression. *Biological Psychiatry*, *53*(8), 649–659.

Fava, M., & Davidson, K. G. (1996). Definition and epidemiology of treatment-resistant depression. *Psychiatric Clinics of North America*, *19*(2), 179–200.

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*(2), 177–184.

Gaynes, B.N., Asher, G., Gartlehner, G., Hoffman, V., Green, J., Boland, J., … Bann, C. 2018. Definition of treatment-resistant depression in the Medicare population. Technology Assessment Program–Technical Assessment. Project ID: PSYT0816 (Prepared by RTI–UNC Evidence-Based Practice Center under Contract No. HHSA2902015 00011I_HHSA29032006T), Rockville, MD.

Gaynes, B. N., Lloyd, S. W., Lux, L., Gartlehner, G., Hansen, R. A., Brode, S., … Lohr, K. N. (2014). Repetitive transcranial magnetic stimulation for treatment-resistant depression: A systematic review and meta-analysis. *Journal of Clinical Psychiatry*, *75*(5), 477–489.

Gibson, T. B., Jing, Y., Smith Carls, G., Kim, E., Bagalman, J. E., Burton, W. N., … Goetzel, R. Z. (2010). Cost burden of treatment resistance in patients with depression. *The American Journal of Managed Care*, *16*(5), 370–377.

Hidalgo-Mazzei, D., Berk, M., Cipriani, A., Cleare, A. J., Florio, A. D., Dietch, D., … Stokes, P. R. A. (2019). Treatment-resistant and multi-therapy-resistant criteria for bipolar depression: Consensus definition. *British Journal of Psychiatry*, *214*(1), 27–35.

Ijaz, S., Davies, P., Williams, C. J., Kessler, D., Lewis, G., & Wiles, N. (2018). Psychological therapies for treatment-resistant depression in adults. *Cochrane Database of Systematic Reviews*, *5*, 010558.

Institute for Clinical Systems Improvement (2016). *Depression, Adult in Primary Care*. Bloomington, MN: Institute for Clinical Systems Improvement.

Ivanova, J. I., Birnbaum, H. G., Kidolezi, Y., Subramanian, G., Khan, S. A., & Stensland, M. D. (2010). Direct and indirect costs of employees with treatment-resistant and non-treatment-resistant major depressive disorder. *Current Medical Research and Opinion*, *26*(10), 2475–2484.

Kautzky, A., Dold, M., Bartova, L., Spies, M., Kranz, G. S., Souery, D., … Kasper, S. (2019). Clinical factors predicting treatment-resistant depression: Affirmative results from the European multicenter study. *Acta Psychiatrica Scandinavica*, *139*(1), 78–88.

Kessler, R. C., Berglund, P., Demler, O., Jin, R., Koretz, D., Merikangas, K. R., … Wang, P. S. (2003). The epidemiology of major depressive disorder: Results from the National Comorbidity Survey Replication (NCS-R). *Journal of the American Medical Association*, *289*(23), 3095–3105.

Lam, R. W., Kennedy, S. H., Grigoriadis, S., McIntyre, R. S., Milev, R., Ramasubbu, R., … Ravindran, A. V., Canadian Network for M and others (2009). Canadian Network for Mood and Anxiety Treatments (CANMAT) Clinical guidelines for the management of major depressive disorder in adults. *Journal of Affective Disorders*, *117*(Suppl 1), S26–S43.

Malhi, G. S., Bassett, D., Boyce, P., Bryant, R., Fitzgerald, P. B., Fritz, K., … Singh, A. B. (2015). Royal Australian and New Zealand College of Psychiatrists clinical practice guidelines for mood disorders. *Australian and New Zealand Journal of Psychiatry*, *49*(12), 1087–1206.

National Institute for Health and Clinical Excellence (NICE) (2018). *Depression. The treatment and management of depression in adults (updated edition)*. Leicester, UK: British Psychological Society.

National Institute of Mental Health (2017). *Prevalence of major depressive episodes among adults*. Bethesda, MD: National Institute of Mental Health.

Naumann, J., Grebe, J., Kaifel, S., Weinert, T., Sadaghiani, C., & Huber, R. (2017). Effects of hyperthermic baths on depression, sleep and heart rate variability in patients with depressive disorder: A randomized clinical pilot trial. *BMC Complementary and Alternative Medicine*, *17*(1), 172.

NeuroStar Advanced Therapy (2017). *About NeuroStar® TMS Therapy*. Malvern, PA: NeuroStar Advanced Therapy.

New England Comparative Effectiveness Public Advisory Council (2012). Treatment options for treatment-resistant, *Depression: Repetitive transcranial magnetic stimulation*. Boston, MA: Institute for Clinical and Economic Review.

Ng, C. H., Kato, T., Han, C., Wang, G., Trivedi, M., Ramesh, V., … Kasper, S. (2019). Definition of treatment-resistant depression—Asia Pacific perspectives. *Journal of Affective Disorders*, *245*, 626–636.

Oquendo, M. A., Baca-Garcia, E., Kartachov, A., Khait, V., Campbell, C. E., Richards, M., … Mann, J. J. (2003). A computer algorithm for calculating the adequacy of antidepressant treatment in unipolar and bipolar depression. *Journal of Clinical Psychiatry*, *64*(7), 825–833.

Papadimitropoulou, K., Vossen, C., Karabis, A., Donatti, C., & Kubitz, N. (2016). Comparative efficacy and tolerability of pharmacological and somatic interventions in adult patients with treatment-resistant depression: A systematic review and network meta-analysis. *Current Medical Research and Opinion*, 1–27.

Perlis, R. H., Ostacher, M. J., Patel, J. K., Marangell, L. B., Zhang, H., Wisniewski, S. R., … Thase, M. E. (2006). Predictors of recurrence in bipolar disorder: Primary outcomes from the Systematic Treatment Enhancement Program for Bipolar Disorder (STEP-BD). *American Journal of Psychiatry*, *163*(2), 217–224.

Ruhé, H. G., van Rooijen, G., Spijker, J., Peeters, F. P. M. L., & Schene, A. H. (2012). Staging methods for treatment resistant depression. A systematic review. *Journal of Affective Disorders*, *137*(1–3), 35–45.

Rush, A. J., Trivedi, M. H., Wisniewski, S. R., Nierenberg, A. A., Stewart, J. W., Warden, D., … Fava, M. (2006). Acute and longer-term outcomes in depressed outpatients requiring one or several treatment steps: A STAR*D report. *American Journal of Psychiatry*, *163*(11), 1905–1917.

Sackeim, H. A. (2001). The definition and meaning of treatment-resistant depression. *Journal of Clinical Psychiatry*, *62*(Suppl 16), 10–17.

Schosser, A., Serretti, A., Souery, D., Mendlewicz, J., Zohar, J., Montgomery, S., & Kasper, S. (2012). European Group for the Study of Resistant Depression (GSRD)–where have we gone so far: Review of clinical and genetic findings. *European Neuropsychopharmacology*, *22*(7), 453–468.

Silverstein, W. K., Noda, Y., Barr, M. S., Vila-Rodriguez, F., Rajji, T. K., Fitzgerald, P. B., … Blumberger, D. M. (2015). Neurobiological predictors of response to dorsolateral prefrontal cortex repetitive transcranial magnetic stimulation in depression: A systematic review. *Depression and Anxiety*, *32*(12), 871–891.

Singh, J., Singh, A., Kar, S. K., & Pahuja, E. (2017). Early augmentation response with low-frequency repetitive transcranial magnetic stimulation in treatment-resistant depression. *Clinical Psychopharmacology and Neuroscience*, *15*(2), 197–198.

Souery, D., Amsterdam, J., de Montigny, C., Lecrubier, Y., Montgomery, S., Lipp, O., … Mendlewicz, J. (1999). Treatment-resistant depression: Methodological overview and operational criteria. *European Neuropsychopharmacology*, *9*(1-2), 83–91.

Souery, D., Oswald, P., Massat, I., Bailer, U., Bollen, J., Demyttenaere, K., … Mendlewicz, J. (2007). Clinical factors associated with treatment resistance in major depressive disorder: Results from a European multicenter study. *Journal of Clinical Psychiatry*, *68*(7), 1062–1070.

Strawbridge, R., Carter, B., Marwood, L., Bandelow, B., Tsapekos, D., Nikolova, V. L., … Young, A. H. (2019). Augmentation therapies for treatment-resistant depression: Systematic review and meta-analysis. *British Journal of Psychiatry*, *214*(1), 42–51.

Thase, M. E., & Rush, A. J. (1997). When at first you don't succeed: Sequential strategies for antidepressant nonresponders. *Journal of Clinical Psychiatry*, *58*(Suppl 13), 23–29.

Thomas, S. P., Nandhra, H. S., & Jayaraman, A. (2010). Systematic review of lamotrigine augmentation of treatment resistant unipolar depression (TRD). *Journal of Mental Health*, *19*(2), 168–175.

Trevino, K., McClintock, S. M., McDonald Fischer, N., Vora, A., & Husain, M. M. (2014). Defining treatment-resistant depression: A comprehensive review of the literature. *Annals of Clinical Psychiatry*, *26*(3), 222–232.

Trivedi, M. H., Rush, A. J., Wisniewski, S. R., Nierenberg, A. A., Warden, D., Ritz, L., … Fava, M. (2006). Evaluation of outcomes with citalopram for depression using measurement-based care in STAR*D: Implications for clinical practice. *American Journal of Psychiatry*, *163*(1), 28–40.

Trivedi, R. B., Nieuwsma, J. A., & Williams, J. W. (2011). Examination of the utility of psychotherapy for patients with treatment resistant depression: A systematic review. *Journal of General Internal Medicine*, *26*, 643–650.

United Nations Development Programme (2015). *Human development report 2015: Work for human development*. New York, NY: United Nations Development Programme.

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*(6), 629–640.

Washington State Health Care Authority (2014). Nonpharmacologic treatments for treatment-resistant depression. https://www.hca.wa.gov/assets/program/TRD_final_report_022114[1].pdf

Yoon, S., Jeon, S. W., Ko, Y. H., Patkar, A. A., Masand, P. S., Pae, C. U., & Han, C. (2017). Adjunctive brexpiprazole as a novel effective strategy for treating major depressive disorder: A systematic review and meta-analysis. *Journal of Clinical Psychopharmacology*, *37*(1), 46–53.

Zhang, Y. Q., Zhu, D., Zhou, X. Y., Liu, Y. Y., Qin, B., Ren, G. P., & Xie, P. (2015). Bilateral repetitive transcranial magnetic stimulation for treatment-resistant depression: A systematic review and meta-analysis of randomized controlled trials. *Brazilian Journal of Medical and Biological Research*, *48*(3), 198–206.

SUPPORTING INFORMATION

Additional supporting information may be found online in the Supporting Information section.

How to cite this article: Gaynes BN, Lux L, Gartlehner G, et al. Defining treatment-resistant depression. *Depress Anxiety*. 2019;1–12. https://doi.org/10.1002/da.22968