Special Article - Depression Disorders & Treatment

A Review of the Nature and Impact of Exclusion Criteria in Depression Treatment Outcome Research

Halvorson MA[1]* and Humphreys K[2]

[1]Department of Veterans Affairs Palo Alto Health Care System, Center for Innovation to Implementation, USA

[2]Department of Psychiatry & Behavioral Sciences, Stanford University, USA

*Corresponding author: Max A. Halvorson, VA Palo Alto Health Care System, Center for Innovation to Implementation (152-MPD), 795 Willow Rd, Menlo Park, CA 94025, USA, Tel: (650)493-5000 x20514; Email: Max. Halvorson@va.gov

Abstract

Background: Depression treatment outcome research is typically designed to maximize internal validity. However, many studies utilize extensive exclusion criteria that reduce the extent to which study samples resemble clinical populations. A better understanding of exclusion criteria and their use in depression treatment outcome research is necessary to generalize accurately from studies. This review identifies the most commonly-used exclusion criteria, the proportion of potential participants excluded, and differences between participant patient samples and patient samples excluded from studies.

Methods: Eighteen studies of exclusion in depression treatment research were identified through PubMed and reviewed by both authors.

Results: A typical study of depression utilizes approximately 8-11 exclusion criteria and excludes between 75% and 85% of depressed individuals. Insufficient depression severity and comorbid Axis I disorders excluded the most potential participants. Excluded individuals tended to be younger and male, and to have poorer mental health status in the form of longer depressive episodes and more psychiatric comorbidities. Treatment response and remission were generally worse for excluded patients than enrolled subjects. Conclusion: Most depressed individuals would be excluded from a typical clinical trial of depression. These excluded individuals differ significantly in baseline characteristics and outcomes from study participants. More inclusive samples and more thorough reporting of exclusion criteria are needed to safely generalize from depression treatment outcome studies.

Keywords: Depression; Clinical trials; Treatment; Antidepressants; CBT/cognitive behavior therapy

Background

Methodologically sound clinical trials of pharmacotherapy and psychotherapy for depression can inform evidence-based care for this prevalent and disabling disorder [e.g., 1,2]. In such trials, depression treatment researchers typically strive to maximize sample homogeneity by excluding patients with particular characteristics (e.g., suicidal impulses, alcohol dependence). The use of exclusion criteria can protect the safety of potential participants and can also sometimes increase ability to detect significant differences in treatment outcome with in small samples, but may have the unintended consequence of lowering generalizability [3,4]. For example, about 4 in 5 patients with schizophrenia are excluded from schizophrenia treatment research [5] and an even higher proportion of patients with cardiovascular disease are excluded from treatment research focused on their disorder [6]. When such large proportions of individuals are not studied in a clinical research area, it becomes less plausible that the results of treatment research can be generalized to front-line healthcare.

To our knowledge, no systematic review has been conducted in the depression field regarding whether exclusion criteria create treatment research samples that differ significantly from real-world patient samples in their clinical characteristics and outcomes.

Accordingly, this review summarizes the evidence base on the use and impact of exclusion criteria in clinical research on depression. We first present the prevalence of specific exclusion criteria employed in the literature, followed by a summary of reported rates of overall exclusion. Next, we present percentages of participants excluded by each individual criterion. Finally, we summarize evidence linking exclusions to sample characteristics and outcomes.

Methods

The Cross-Disease Review of Exclusion Across Medicine (CREAM) project is a structured literature review of studies of exclusion criteria and their impact across a range of disciplines (e.g., oncology, cardiology, psychiatry). Methods are described in detail elsewhere [5], but to summarize: literature was identified by conducting English-language searches in PubMed on the following terms: "Eligibility criteria and generalizability" (anywhere in paper), "exclusion criteria and generalizability" (anywhere in paper), "exclusion criteria" (in title of paper) and "eligibility criteria" (in title of paper). To be considered relevant, studies had to analyze data on [1] The prevalence and nature of exclusion criteria in a particular field, and/or [2] The impact of exclusion criteria on sample representatives or study results. Reference lists of all studies identified in the search were themselves scanned for more potential studies.

Table 1: Prevalence of exclusion criteria.

Exclusion criterion	% of studies using criterion		
	Posternak et al. (N of studies=31) [13]	Zimmerman et al. (N of studies=39) [14]	Van der Lem et al. (N of studies=20) [12]
Depression severity too low	96.7	92.3	80
Suicidal ideation	-	66.7	40
Episode duration too short	41.9	48.6 (*either* too long *or* too short)	-
Episode duration too long	12.9		-
Substance use disorder history (current or recent)	83.9	82.1	85
Psychotic features	-	87.1	90
Mania or Bipolar Disorder	-	48.7	-
Anxiety Disorder	35.5	-	45
Dysthymia	19.4	28.2	-
Any Axis I Disorder	59	-	-
Borderline Personality Disorder	-	20.5	95
Any personality disorder	16.1	-	60
Response to treatment during lead-in period	54.8	-	-
Prior non response to treatment	48.4	-	-
Recent treatment with other antidepressants or electroconvulsive therapy	-	-	70
Previous psychotherapy	-	-	40
Concomitant therapy	-	-	50
Medical contraindication	-	-	45
Cognitive disorders	-	-	55
Somatization disorders	-	-	55
Other psychiatric comorbidity (eating disorder, obsessive-compulsive disorder)	-	-	25

Results

Our search terms yielded three reviews of the prevalence of specific exclusion criteria in clinical trials of depression treatment and 15 empirical examinations of study exclusion in depression treatment research. Three of the studies focused on exclusion criteria in psychotherapy efficacy trials and the remaining 12 studies focused on exclusion in antidepressant efficacy trials.

Prevalence of exclusion criteria in depression treatment trials

Exclusion criteria are utilized in virtually all clinical trials; however, the number of criteria used varies greatly across studies. For example, in the five retrospective studies which included information on specific exclusion criteria used in the original trial, between 3 and 21 criteria were reported [7-11]. In a study of 20 psychotherapy efficacy trials by van der Lem [12], the authors identified 38 unique exclusion criteria which they categorized into 15 groups. Eight of these criteria appeared in 50% or more of the studies and the average study used eight of these 15 criteria.

We identified three reviews of the depression literature which sought to describe the prevalence of common exclusion criteria in depression treatment studies (Table 1). Posternak, Zimmerman [13] characterized rates of exclusion criteria use in all 31 antidepressant efficacy trials published between 1994 and 1998 in 5 major psychiatry journals. The authors excluded studies which focused on a specific subgroup of patients. Zimmerman, Chelminski [14] reported rates of exclusion in all 39 antidepressant efficacy trials in 5 major psychiatry journals published from 1994 to 2000. It should be noted that the sample of studies in (Posternak, Zimmerman [13]) is a subset of the sample of studies in Zimmerman, Chelminski [14]. However, as the two reviews reported rates of use for different criteria, each provides unique data. A third study by van der Lem, de Wever [12] presented rates of exclusion criteria for 20 trials of psychotherapy in adults found in a literature search of PubMed and PsycInfo.

Certain criteria are found in most depression treatment trials. Insufficient depression severity (depression scores in the "mild" range as measured by the Beck Depression Inventory [BDI], the Hamilton Depression Rating Scale [HDRS], or the Montgomery-Asberg

Table 2: Percent excluded by criterion and overall exclusion rates.

	N	Any Axis I	Anxiety	Psychosis	Bipolar	Other Axis I	Personality Disorder	Border-line	Dysthymia	Substance Use	Suicidality	Low Dep. Score	Short Episode	Long Episode	Other Meds	Physical Issue	ECT	Prior Non-Response	Total % Excluded
Blanco, 2008 [19]	3119	47.4	-	2.4	17.4	-	-	-	16.0	8.8	8.9	-	40.3		-	-	-	-	75.8
Haberfellner, 2000 [7]	216	24.0	-	-	-	-	-	-	-	5.6	4.2	91.7	-	-	-	-	-	-	100.0
Keitner, 2002 [8]	186	-	-	-	1.2	-	-	-	-	9.4	-	8.2	-	-	7.8	-	-	19.9	85.5
Partonen, 1996 [9]	612	-	-	11.0	-	-	-	-	-	17.0	9.0	-	-	-	15.0	14.0	4.0	-	62.3
Schindler, 2011 [23]	338	-	-	0.7	-	-	-	5.9	-	12.9	2.0	23.7	-	-	-	-	-	-	24.0
Seemuller, 2010 [20]	971	14.2	-	8.0	6.6	-	-	1.9	5.6	8.3	12.0	27.0	5.0	7.2	-	-	-	-	68.8
Sullivan, 1994 [10]	95	-	-	-	-	-	-	18.9	-	28.4	-	-	-	-	-	-	-	-	51.6-61.1
van der Lem, 2011 [21]	1653	62.8	-	1.9	3.4	-	31.6-61.6	0.2-7.0	8.5	8.6	15.2	27.2-41.6	-	-	-	-	-	-	75.5-83.0
van der Lem, 2012 [12]	598	-	-	6.5		-	-	-	-	7.3	-	21.9	-	-	45.0	-	-	-	-
Westen, 2001 [24]	1108	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	68.1
Wisniewski, 2009 [3]	2855	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	77.8
Yastrubetskaya, 1997 [11]	188	56.3	-	10.1	8.5	-	-	-	4.3	-	-	-	-	-	95.0	-	11.1	-	95.7
Zetin, 2007 [22]	348	-	46.6	1.1	19.8	8.5	-	0.0	9.3	5.3	17.8	59.1	0.0	23.1	-	-	-	-	91.4
Zimmerman, 2002 [18]	346	-	64.3	7.0	9.0	5.0	-	2.6	2.7	12.7	1.7	54.3	2.4	19.4	-	-	-	-	91.6
Zimmerman, 2004 [14]	596	-	55.5	5.7	9.9	8.0	-	10.1	8.9	8.9	6.6	32.4-47.0	6.4	32.0	-	-	-	-	65.8 (mean)

Depression Rating Scale [MADRS]; [15-17]) was the most commonly used exclusion criterion, followed by substance use disorder history, presence of psychotic symptoms, and borderline personality disorder. These four criteria appeared in over 80% of depression efficacy studies examined in at least one of the three reviews presented in Table 1. Suicidal ideation was also a common exclusion criterion, appearing in 66.7% of 39 studies in Zimmerman et al. [14] review. Other Axis I diagnoses such as bipolar disorder (48.7% of studies in Zimmerman et al.), anxiety disorders (35.5% in Posternak et al. and 45.0% in van der Lem et al.), and dysthymia (19.4% in Posternak et al. and 28.2% in Zimmerman et al.) too were commonly used to screen patients out of study samples. Zimmerman's and van der Lem's reviews suggest that a typical study of depression makes use of approximately 8-11 exclusion criteria.

Based on their analysis of 31 depression treatment studies published in five leading psychiatry journals between 1994 and 1998, Zimmerman, Mattia [18] identified a set of 11 commonly used exclusion criteria: history of mania or hypomania, psychotic features during the current depressive episode, significant risk of suicide, comorbid anxiety disorders, alcohol or drug use disorder in the past 6 months, insufficient depression severity, dysthymic disorder, a depressive episode shorter than 4 weeks, a depressive episode longer than 2 years, any comorbid axis I disorder, and borderline personality disorder. These 11criteria have been widely utilized by other investigators to estimate the impact of exclusion criteria in depression treatment research [14,19-22].

Proportion of potential participants excluded by individual criteria

Individual criteria excluded potential participants to widely varying extents (Table 2). In the studies of exclusion criteria examined, a few individual criteria excluded over 50% of potential participants in at least one study: insufficient depression severity, depressive episode duration that was either too short or too long, comorbid Axis I disorders (including anxiety or mood disorders), and personality psychopathology. Another set of criteria excluded over 15% of potential participants in at least one study: bipolar disorder or history of mania or hypomania, borderline personality disorder, underlying dysthymia, current or prior substance use disorder or dependence, suicidal ideation, use of another antidepressant, and prior nonresponsive to treatment.

Proportion of potential participants excluded across all criteria

The final column of Table 2 summarizes overall exclusion rates for 14 studies of exclusion in depression treatment. The majority (nine) of these studies simulated trial exclusion by taking widely-inclusive patient pools, applying common exclusion criteria to the pre-exclusion sample, and reporting overall percentages of participants excluded and percentages of participants excluded for each criterion [3,10,14,18-23]. Another set of five studies retrospectively studied exclusion in trials which had already been completed [7-9,11,24].

As mentioned, much of the literature on exclusion criteria in depression treatment research examines exclusion criteria identified by Zimmerman, Mattia [18]. These researchers found that 317 of 346 (91.6%) depressed outpatients in a clinical practice would have been excluded by this set of exclusion criteria. In more recent work, Zetin and Hoepner [22] replicated Zimmerman's exclusion methodology with a separate clinical practice sample. Though relative exclusion rates for individual criteria differed, Zetin and Hoepner found a strikingly similar proportion (91.4%) of their patient sample was excluded by at least one criterion from a sample of 348 consecutive clinical practice patients. In both of these studies, the three criteria excluding the most patients were comorbid anxiety disorder (64.3% in Zimmerman, 46.6% in Zetin), low depression severity score (54.3% in Zimmerman, 59.1% in Zetin), and a depressive episode duration longer than two years (19.4% in Zimmerman, 23.1% in Zetin).

Several researchers applied Zimmerman's set of standard exclusion criteria to currently depressed individuals identified in epidemiological surveys. Blanco, Olfson [19] did so in the National Epidemiologic Survey for Alcohol and Related Conditions (N=3119) and found that 75.8% of 3,119 depressed individuals would be excluded from treatment research. Axis I comorbidities and an episode duration either less than four weeks or more than two years were the primary excluders. Applying Zimmerman's criteria only to depression treatment-seekers (N=1359) in the same sample, Blanco, Olfson [19] found an exclusion rate of 66.9%. Seemuller, Moller [20] utilized Zimmerman's criteria to determine that 68.8% of 971 patients in German psychiatric university and district hospitals would be excluded from depression treatment research. In a naturalistic study in the Netherlands involving routine outcome monitoring by research nurses, van der Lem, van der Wee [21] simulated rates of exclusion among 1,653 patients with depression using Zimmerman's criteria. The research team found that 75.5%-83.0% of patients were excluded, depending on the equation used to convert MADRS scores to Zimmerman's BDI cutoffs.

Wisniewski, Rush [3] examined typical phase III clinical trial exclusion criteria in data from the Sequenced Treatment Alternatives to Relieve Depression (STAR*D) project. The specific set of typical exclusion criteria was selected by a group of the authors with extensive background in conducting clinical trials. Many, but not all (e.g., suicidal ideation was not examined) of Zimmerman's exclusion criteria were analyzed, but some were operationalized differently (e.g., substance use disorder was only an exclusion if detoxification was required). Of the 2,855 participants in the STAR*D trial, 77.8% would have been excluded from a typical phase III clinical trial, and even this estimate was likely downwardly biased since participants with a HAM-D score of less than 14 were removed from the sample before the exclusion analyses. Even so, Wisniewski's results were similar to those of Zimmerman and others.

In a small-scale study of depression treatment response, Sullivan and Joyce [10] examined exclusion under two sets of criteria: the more restrictive NIMH Treatment of Depression Collaborative Research Program (TDCRP; [25,26]) criteria and the more inclusive Maintenance Therapies in Recurrent Depression Protocol (MTRDP; [27,28]) criteria. The authors found an exclusion rate for the 95 patients of 61.1% under the TDCRP criteria and a rate of 51.6% under the MTRDP criteria. These results are likely biased downwards, however, as patients with a past manic episode, a medical contraindication, current moderate to severe drug or alcohol dependence, or concurrent pharmacotherapy were excluded from the sample prior to the assessment of the impact of the MTRDP and TDCRP criteria.

Zimmerman, Chelminski [14] applied sets of exclusion criteria from 39 antidepressant efficacy trials published in five major psychiatry journals between 1994 and 2000 to a sample of 596 depressed outpatients, finding a mean exclusion rate of 65.8%. Exclusion rates for studies in the sample ranged from 0.0% to 95.0%, with more than 70% excluded in approximately half of the studies. Exclusion rates for specific criteria mirrored closely results from Zimmerman's [18] earlier work with the most commonly met exclusion criteria being comorbid anxiety, low depression score, and long depressive episode duration.

Schindler, Hiller [23] reported that 24.0% of depressed patients (N=338) at a university outpatient clinic would be excluded under a set of exclusion criteria. However, the set of exclusion criteria used by the authors was limited to any exclusion criteria used by at least three of five selected RCTs. Episode duration and depressive severity, another Axis I diagnosis, and comorbid personality psychopathology, all of which would have excluded numerous patients, were not considered as exclusion criteria in Schindler's study. These factors led to a lower estimate of exclusion than in otherwise similar studies. Moreover, nearly all of Schindler's reported exclusions could be attributed to a depression score which reflected too mild a form of depression.

Five studies retrospectively examined exclusion in clinical trials using study records. These studies provide primary data on the extent of exclusion in clinical trials, but do not use a standard set of exclusion criteria and thus their overall exclusion rates are difficult to compare to other studies of exclusion. Note also that because the number of excluded individuals is not always carefully tracked, published exclusion rates are probably generally conservative [29,30].

In a meta-analysis by Westen and Morrison [24] including 12 studies of depression, 68.1% of patients were excluded. A second study by Keitner, Posternak [8] retrospectively examined rates of exclusion for two clinical trials between 1997 and 2002, finding that 85.5% of interested depressed individuals (N=186) were excluded. The exclusion criteria used in the RCTs examined by Keitner overlapped with Zimmerman's, but did not include psychiatric comorbidities and were operationalized differently in some cases. In a third retrospective study of exclusion in a double-blind comparative trial of two antidepressants in Finland, Partonen [9] and colleagues found that 62.3% of the 612 potential participants interviewed were excluded from the trial. As comorbid psychiatric disorders did not exclude patients in this study, this estimate of overall exclusion may be slightly low. In recruitment of a sample of older adults for a phase III clinical trial of an antidepressant, Yastrubetskaya, Chiu [11] tracked exclusion rates for a group of older adults (N=188) and found that 95.7% were excluded under a set of study-specific criteria. Yastrubetskaya's work involves only older adults and used a broad set of exclusion criteria; thus, more patients are excluded than one might expect from a typical trial of adults. A final study by Haberfellner [7] tracked recruitment efforts at an outpatient psychiatric practice for a study of individuals presenting with depressive symptoms. The 21 criteria used screened out a remarkable 100% of potential participants in a recruitment sample of 216 consecutive patients and the average number of exclusion criteria met by a potential participant was 3.4.

Best Estimates of Overall Exclusion Rates. The literature on exclusion criteria in depression suggests that a strong majority of individuals with depression would be excluded from most trials on depression treatment efficacy. One study with a design that would downwardly bias its findings was an outlier with an exclusion rate of 24.0%. Exclusion rates for the other studies ranged from 51% to 100%. Thirteen studies excluded over 60% of potential participants, nine excluded over 70%, five excluded over 80%, and three excluded over 90%.

In the four studies with the largest sample sizes (N > 1,000; 3, 19, 21, 24), total exclusion rates ranged from 68% to 83%. The studies conducted in clinics [18,22], in which face-to-face contact with

Table 3: Differences in sample composition between patients who would be excluded by typical exclusion criteria and patients who would be included.

Study	N	Demographic Differences	Clinical Differences
Partonen, 1996 [9]	612	Excluded were younger (47.0 vs. 44.0 years), more likely male (44.0% vs. 35.0%), more often unmarried (25.0% vs. 16.0%)	Excluded were more likely to have comorbid psychiatric disorders (32.0% vs. 14.0%), more often had a history of MDD (59.0% vs. 49.0%), lower BDI score at baseline (score not reported, p = .002) and less likely to have a current depressive episode (49.0% vs. 61.0%)
Schindler, 2011 [23]	338	No demographic differences	Excluded were less likely to have anxiety disorders (22.2% vs. 35.4%)
Seemuller, 2010 [20]	971	Excluded were younger (44.5 vs. 46.3 years), more likely to be treated in a university hospital vs. a district hospital (77.1% vs. 69.6%)	Excluded were more likely to have comorbid psychiatric disorders (42.2% vs. 10.3%)
Sullivan, 1994 [10]	95	Excluded were younger (27.7-28.1 vs. 35.2-36.3 years), more men (55.0-57.0% vs. 35.0-37.0%)	Excluded had more comorbid psychiatric disorders (2.9-3.2 vs. 1.8-1.9)
Wisniewski, 2009 [3]	2855	Excluded were older (41.5 vs. 38.3 years), less educated (13.2 vs. 14.4 years of education), poorer (monthly household income of $2,163 vs. $3,050), more likely Black (19.2% vs. 11.7%), more likely Hispanic (14.4% vs. 8.5%), more likely unemployed (40.3% vs. 30.8%), less likely to have private insurance (48.3% vs. 61.4%)	Excluded were more likely to have a past suicide attempt (18.7% vs. 15.1%), had longer illness duration(16.1 vs. 13.5 years)

clinicians and clinical staff likely increased accuracy of diagnosis and classification of exclusions, found exclusion rates around 90%. Zimmerman, Chelminski [14] applied various sets of exclusion criteria used in published trials to a single dataset and found a mean exclusion rate of 68%. The remaining studies included either few exclusion criteria [23] or small samples (N < 100; 10). The three studies which retrospectively examined exclusion rates from individual trials [8,9,24] found mean exclusion rates of 86%, 62%, and 68%. As these studies used a smaller subset of common exclusion criteria, they likely provide a downwardly-biased estimate of the likelihood that a depressed patient would meet a commonly-used exclusion criterion. Thus, weighting the large-scale studies (which excluded around 75%) and the clinic studies (around 90%) most strongly, and taking into account the likely downwardly-biased estimates from retrospective analyses (around 70%), our best estimate of the percentage of depressed patients who would be excluded from most clinical trials of depression treatment lies between 75% and 85%.

Effects of exclusion criteria on sample representativeness

We identified five studies which compared demographic and clinical characteristics of research samples to excluded patients and reported significant differences (Table 3). Partonen, Sihvo [9] compared excluded patients from participants in a clinical trial, finding that their group of excluded patients contained more men than did the sample of participants, and that those excluded were younger and more likely to be unmarried. Men were more likely than women to be excluded due to suicidal ideation or chronic alcohol or drug misuse. Excluded participants were also more likely to have comorbid psychiatric disorders and a history of major depression but were less depressed on the BDI and less likely to have a current major depressive episode.

Four studies examined demographic and clinical differences between those who would have been excluded from a clinical trial and those who would have been participants. In Schindler et al. [23] analyses, the authors found no demographic differences between potentially included and potentially excluded participants. Potentially excluded participants were less likely to have comorbid anxiety disorders. Seemuller and colleagues [20] reported that excluded participants were younger, and more likely to be treated in a university hospital vs. a district hospital. Rates of comorbid psychiatric disorders were higher for patient who would have been excluded. Sullivan and Joyce [10] found that patients who would

have been excluded were younger than included patients and more often male. They were also more likely to have comorbid psychiatric disorders. In work by Wisniewski, Rush [3], those excluded were older, less educated, poorer, more likely to be Black or Hispanic, more likely to be unemployed, and less likely to have private insurance. Clinically, those excluded were more likely to have had a past suicide attempt and typically had longer depressive episodes.

Four of the five studies found demographic differences between excluded and included patients. Though the specific pattern of differences was not consistent, excluded individuals were generally younger and more likely to be male than included participants. One study each found excluded patients to be more often unmarried, more likely to be treated in a university hospital (vs. a district hospital), less educated, poorer, more often Black or Hispanic, and less likely to have private insurance. All five of the studies found differences in the clinical profiles of excluded individuals and included individuals. Excluded patients were more likely to have comorbid psychiatric disorders in three of the five studies, although one study found lower levels of anxiety disorders in excluded patients. One study found a higher proportion of excluded patients with a prior major depressive episode and another found longer average illness duration amongst excluded patients. Furthermore, excluded patients were more likely to have had a suicide attempt in one study. In contrast, one study found that excluded patients had lower BDI scores than included patients.

Effects of exclusion criteria on sample treatment outcomes

We identified five studies comparing depression outcomes for patients who would have been excluded under standard exclusion criteria to those of included individuals (Table 4). The studies focused on trials of clinically depressed individuals which included minimal exclusion criteria (psychiatric comorbidities were not exclusion criteria), comparing patients who met common exclusion criteria to those who did not.

Mental health outcomes and treatment response were generally worse and in no cases better for excluded individuals. Seemuller, Moller [20] observed slightly poorer overall functioning at follow-up for excluded individuals, as evidenced by a lower Global Assessment of Functioning (GAF) score. Sullivan and Joyce [10] found no differences in outcome between excluded and included fewer than three sets of exclusion criteria, but poorer treatment response as

Table 4: Differences in outcomes between patients who would be excluded by typical exclusion criteria and patients who would be included.

Study	Comparison	Outcome Differences
Seemuller, 2010 [20]	Excluded *vs.* included	Excluded had lower Global Assessment of Functioning score (69.3 *vs.* 71.4) at follow-up
Sullivan, 1994 [10]	Excluded *vs.* included	Excluded had less treatment response (HAM-D improvement; 56.5% *vs.* 71.0%) under TDCRP criteria
van der Lem, 2011 [21]	Excluded *vs.* Included	No differences in outcomes (response or remission)
van der Lem, 2012 [12]	Excluded *vs.* full sample	Excluding those with insufficient depression severity led to fewer in the sample being remitted (OR=0.53). Excluding those who had previously received other medications at baseline led to a marginally greater number in the sample who responded (OR=1.47) or remitted (OR=1.53) and a greater number responding to treatment.
Wisniewski, 2009 [3]	Excluded *vs.* included	Excluded had lower rates of treatment response (39.1% *vs.* 51.6%), remission (24.7% *vs.* 34.4%), lower self-reported maximum side effect intensity and burden, higher likelihood of serious adverse events (4.5% *vs.* 2.4%) and serious psychiatric adverse events (2.3% *vs.* 0.9%).

defined by reduction on the Hamilton Depression Rating Scale (HDRS) for those excluded under a fourth set, the NIMH Treatment of Depression Collaborative Research Program (TDCRP) criteria. Wisniewski, Rush [3] noted that excluded participants experienced lower rates of treatment response and remission, higher subjective maximum side effect intensity and burden, and a higher likelihood of serious adverse events and serious psychiatric adverse events.

Although van der Lem, de Wever [12] compared the total (both included and excluded) sample to the included sample, making their statistical results difficult to interpret, the authors found that excluding those with low depression severity led to fewer in the sample being remitted. They also found that excluding those who had previously received medication or ECT for depression led to a marginally greater number in the sample being remitted and a marginally greater number responding to treatment, suggesting that those excluded due to prior depression treatment remitted at a lower rate. A final study by van der Lem, van der Wee [21] examined several outcomes and found no differences in treatment response or remission between excluded and included.

Discussion

Taken together, the studies reviewed establish that exclusion criteria have a substantial impact on depression treatment research study samples. Approximately 75-85% of individuals with depression meet at least one common exclusion criterion for these types of trials. Most studies excluded depressed individuals with a depression severity score reflective of a mild form of depression, those who had experienced suicidal ideation, those who had a history of substance use disorder, those who had psychotic features, those who had a comorbid Axis I disorder (most commonly bipolar disorder or anxiety disorders), and those who had a personality disorder (most commonly borderline personality disorder). Insufficient current depression severity and comorbid Axis I disorders excluded the most potential participants. Based on available evidence, a typical trial utilizes approximately 8-11 exclusion criteria. The evidence, though mixed, suggests that exclusions impose differences between included and excluded samples on demographic (excluded samples were younger and made up of more males) and clinical (excluded samples had more psychiatric comorbidities and longer depressive episode durations) characteristics, and change the outcomes of clinical trials (generally worse and in no case better outcomes for excluded individuals).

Because patients with psychiatric and physical comorbidities – who are especially difficult to treat (e.g., [31-33]) – are disproportionately excluded from clinical trials of depression, clinical trials may overstate how beneficial depression treatment is with unselected patients in front-line care. In our review, studies examining treatment response in patients who would have been excluded from clinical trials found lower rates of treatment response, remission, and general functioning in these populations as compared to trial participants. Some of this effect is undoubtedly due to appropriate exclusion; however, clinicians consistently report that the average patient for which depression treatment is appropriate has substantial comorbidities [8,18,22]. For example, a large-scale epidemiologic survey found rates of current substance use disorders as high as 19.2% among depressed individuals [34]; past or current substance use disorder was an exclusion criteria in nearly every study examined in the current review. As a result, an abundance of depression treatment studies exclude the very patients who are prime candidates for treatment in clinical settings.

The nature and number of exclusion criteria varies substantially from trial to trial. Strikingly, applying unique sets of exclusion criteria from 39 completed clinical trials yielded rates of exclusion ranging from 0.0% to 95.0% in a clinical sample of 1,500 patients [14]. In order to make results replicable, generalizable and understandable, investigators should report exclusion criteria and the extent to which they exclude potential participants as thoroughly as possible. Adherence to the guidelines recommended in the Consolidated Standards of Reporting Trials (CONSORT) 2010 Statement should guide reporting standards for exclusions [35], but for studies of conditions such as depression where comorbidity is common, even greater attention to detail is required in reporting exclusion.

Researchers may worry that more inclusive trial samples might increase sample heterogeneity and thus reduce statistical power. Increased heterogeneity might require larger study samples to detect a main effect of treatment [36]. This can happen, although perhaps surprisingly, exclusion criteria can also have the opposite effect [37]. The advantages of larger, more heterogeneous study samples, however, are twofold: 1) overall results would be more directly generalizable to clinical populations, and 2) subgroup analyses would be possible, allowing researchers to answer open questions about treatment efficacy for comorbid populations. As abundant staff time and resources are devoted to recruiting, screening, and consenting participants in clinical trials, the financial costs of paying a greater number of participants could be offset, at least in part, by a reduction in the number of recruitment efforts per consented participant. Furthermore, multiple research hypotheses regarding patient subgroups, if clearly delineated a priori and appropriately handled, could be answered with a single study [36].

One limitation of the current review is related to the literature itself – most of the studies on exclusion in depression treatment trials focused on pharmacotherapy trials and only a few focused on exclusions in psychotherapy trials. Of the studies included in the present review, only three of fifteen empirical examinations of exclusion focused on psychotherapy. Future research should focus on exclusion rates and reasons for exclusion in psychotherapy trials as they may differ in nature from those used in antidepressant trials. A second potential limitation is the difficulty of converting this particular topic area into effective search terms – as most trials list exclusion information, an unrestricted search for the words "exclusion" or "exclusion criteria" anywhere in a paper would have yielded tens of thousands of mostly irrelevant papers, which was beyond our reviewing resources.

Certain exclusion criteria are undoubtedly necessary for issues of appropriateness of treatment and patient safety, but depression treatment researchers should think as critically about which exclusion criteria they use as they would any other major methodological decision. An exclusion criterion's appearance in a similar previous trial should not necessarily be grounds for utilizing the criterion. By including patient populations who are likely to receive a given treatment in clinical trials, many of whom are multimorbid; trials can more effectively speak to the efficacy of treatment for depression in real clinical practice.

Acknowledgment

References

1. Meldrum ML. A brief history of the randomized controlled trial. From oranges and lemons to the gold standard. Hematol Oncol Clin North Am. 2000; 14: 745-760.

2. Taylor WD, Doraiswamy PM. A systematic review of antidepressant placebo-controlled trials for geriatric depression: limitations of current data and directions for the future. Neuropsychopharmacology. 2004; 29: 2285-2299.

3. Wisniewski SR, Rush AJ, Nierenberg AA, Gaynes BN, Warden D, Luther JF, et al. Can phase III trial results of antidepressant medications be generalized to clinical practice? A STAR*D report. Am J Psychiatry. 2009; 166: 599-607.

4. van der Lem R. Are depression trials generalizable to clinical practice? Rotterdam, Netherlands. Legatron Electronic Publishing. 2013.

5. Humphreys K. A review of the impact of exclusion criteria on the generalizability of schizophrenia treatment research. Clin Schizophr Relat Psychoses. 2014.

6. Hlatky MA, Lee KL, Harrell FE, Califf RM, Pryor DB, Mark DB, et al. Tying clinical research to patient care by use of an observational database. Stat Med. 1984; 3: 375-387.

7. Haberfellner EM. Recruitment of depressive patients for a controlled clinical trial in a psychiatric practice. Pharmacopsychiatry. 2000; 33: 142-144.

8. Keitner GI, Posternak MA, Ryan CE. How many subjects with major depressive disorder meet eligibility requirements of an antidepressant efficacy trial? J Clin Psychiatry. 2003; 64: 1091-1093.

9. Partonen T, Sihvo S, Lonnqvist JK. Patients excluded from an antidepressant efficacy trial. J Clin Psychiatry. 1996; 57: 572-575.

10. Sullivan PF, Joyce PR. Effects of exclusion criteria in depression treatment studies. J Affect Disord. 1994; 32: 21-26.

11. Yastrubetskaya O, Chiu E, O'Connell S. Is good clinical research practice for clinical trials good clinical practice? Int J Geriatr Psychiatry. 1997; 12: 227-231.

12. van der Lem R, de Wever WW, van der Wee NJ, van Veen T, Cuijpers P, Zitman FG. The generalizability of psychotherapy efficacy trials in major depressive disorder: an analysis of the influence of patient selection in efficacy trials on symptom outcome in daily practice. BMC Psychiatry. 2012; 12: 192.

13. Posternak MA, Zimmerman M, Keitner GI, Miller IW. A reevaluation of the exclusion criteria used in antidepressant efficacy trials. Am J Psychiatry. 2002; 159: 191-200.

14. Zimmerman M, Chelminski I, Posternak MA. Exclusion criteria used in antidepressant efficacy trials: consistency across studies and representativeness of samples included. J Nerv Ment Dis. 2004; 192: 87-94.

15. Beck AT, Ward CH, Mendelson M, Mock J, Erbaugh J. An inventory for measuring depression. Arch Gen Psychiatry. 1961; 4: 561-571.

16. Montgomery SA, Asberg M. A new depression scale designed to be sensitive to change. Br J Psychiatry. 1979; 134: 382-389.

17. Hamilton M. A rating scale for depression. J Neurol Neurosurg Psychiatry. 1960; 23: 56-62.

18. Zimmerman M, Mattia JI, Posternak MA. Are subjects in pharmacological treatment trials of depression representative of patients in routine clinical practice? Am J Psychiatry. 2002; 159: 469-473.

19. Blanco C, Olfson M, Goodwin RD, Ogburn E, Liebowitz MR, Nunes EV, et al. Generalizability of clinical trial results for major depression to community samples: results from the National Epidemiologic Survey on Alcohol and Related Conditions. J Clin Psychiatry. 2008; 69: 1276-1280.

20. Seemüller F, Möller HJ, Obermeier M, Adli M, Bauer M, Kronmüller K, et al. Do efficacy and effectiveness samples differ in antidepressant treatment outcome? An analysis of eligibility criteria in randomized controlled trials. J Clin Psychiatry. 2010; 71: 1425-1433.

21. van der Lem R, van der Wee NJ, van Veen T, Zitman FG. The generalizability of antidepressant efficacy trials to routine psychiatric out-patient practice. Psychol Med. 2011; 41: 1353-1363.

22. Zetin M, Hoepner CT. Relevance of exclusion criteria in antidepressant clinical trials: a replication study. J Clin Psychopharmacol. 2007; 27: 295-301.

23. Schindler AC, Hiller W, Witthoft M. Benchmarking of cognitive-behavioral therapy for depression in efficacy and effectiveness studies--how do exclusion criteria affect treatment outcome? Psychotherapy research. Journal of the Society for Psychotherapy Research. 2011; 21: 644-657.

24. Westen D, Morrison K. A multidimensional meta-analysis of treatments for depression, panic, and generalized anxiety disorder: an empirical examination of the status of empirically supported therapies. Journal of consulting and clinical psychology. 2001; 69: 875-899.

25. Elkin I, Parloff MB, Hadley SW, Autry JH. NIMH Treatment of Depression Collaborative Research Program. Background and research plan. Arch Gen Psychiatry. 1985; 42: 305-316.

26. Elkin I, Shea MT, Watkins JT, Imber SD, Sotsky SM, Collins JF, et al. National Institute of Mental Health Treatment of Depression Collaborative Research Program. General effectiveness of treatments. Arch Gen Psychiatry. 1989; 46: 971-982.

27. Frank E, Kupfer DJ, Perel JM, Cornes C, Jarrett DB, Mallinger AG, et al. Three-year outcomes for maintenance therapies in recurrent depression. Arch Gen Psychiatry. 1990; 47: 1093-1099.

28. Kupfer DJ, Frank E, Perel JM, Cornes C, Mallinger AG, Thase ME, et al. Five-year outcome for maintenance therapies in recurrent depression. Arch Gen Psychiatry. 1992; 49: 769-773.

29. Gandhi M, Ameli N, Bacchetti P, Sharp GB, French AL, Young M, et al. Eligibility criteria for HIV clinical trials and generalizability of results: the gap between published reports and study protocols. AIDS. 2005; 19: 1885-1896.

30. Gross CP, Mallory R, Heiat A, Krumholz HM. Reporting the recruitment process in clinical trials: who are these patients and how did they get there? Ann Intern Med. 2002; 137: 10-16.

31. Spinhoven P, Penninx BW, van Hemert AM, de Rooij M, Elzinga BM. Comorbidity of PTSD in anxiety and depressive disorders: prevalence and shared risk factors. Child Abuse Negl. 2014; 38: 1320-1330.

32. Koyuncu A, Ertekin E, Binbay Z, Ozyıldırım I, Yüksel C, Tükel R. The clinical impact of mood disorder comorbidity on social anxiety disorder. Compr Psychiatry. 2014; 55: 363-369.

33. Cully JA, Breland JY, Robertson S, Utech AE, Hundt N, Kunik ME, et al. Behavioral health coaching for rural veterans with diabetes and depression: a patient randomized effectiveness implementation trial. BMC health services research. 2014; 14: 191.

34. Grant BF, Stinson FS, Dawson DA, Chou SP, Dufour MC, Compton W, et al. Prevalence and co-occurrence of substance use disorders and independent mood and anxiety disorders: results from the National Epidemiologic Survey on Alcohol and Related Conditions. Arch Gen Psychiatry. 2004; 61: 807-816.

35. Schulz KF, Altman DG, Moher D. CONSORT Group . CONSORT 2010 statement: updated guidelines for reporting parallel group randomised trials. Int J Surg. 2011; 9: 672-677.

36. Yusuf S, Collins R, Peto R. Why do we need some large, simple randomized trials? Stat Med. 1984; 3: 409-422.

37. Humphreys K, Harris AH, Weingardt KR. Subject eligibility criteria can substantially influence the results of alcohol-treatment outcome research. J Stud Alcohol Drugs. 2008; 69: 757-764.

Citation: Halvorson MA and Humphreys K. A Review of the Nature and Impact of Exclusion Criteria in Depression Treatment Outcome Research. Ann Depress Anxiety. 2015; 2(5): 1058.