

# Generalizability of Findings From Efficacy Trials for Chronic Depression: An Analysis of Eligibility Criteria

Alessa von Wolff, M.A.

Malte Jansen, M.A.

Lars P. Hölzel, Ph.D.

Annika Westphal, B.A.

Martin Härter, M.D., Ph.D.

Levente Kriston, Ph.D.

**Objectives:** Concerns have been raised that the rigorous eligibility criteria used to select patients for randomized controlled trials (RCTs) limit the generalizability of trial findings. The objectives of this study were to identify commonly used eligibility criteria in RCTs for chronic depression, to examine whether these criteria are met by patients with chronic depression who are in routine care, and to identify differences between patients who would and would not meet RCT criteria. **Methods:** Thirteen eligibility criteria were extracted from eight RCTs of combined psychotherapeutic and pharmacological interventions for patients with chronic depression. These criteria were then applied to a sample of patients with chronic depression receiving care in one of ten German hospitals (N=231). Demographic, clinical, and treatment characteristics of those who met the RCT criteria and those who did not were compared in univariate and multivariate analyses. **Results:** Only 25% of the 231 inpatients met all RCT eligibility criteria. Patients were ineligible mainly because of suicide risk, low severity of depression at admission, and concurrent psychiatric or somatic disorders. No statistically significant differences were found between those who met the criteria and those who did not in demographic characteristics, length of inpatient stay, treatment outcome, and efficacy of certain antidepressants, except that slightly more patients meeting RCT criteria received selective serotonin reuptake inhibitors. **Conclusions:** Findings suggest that the generalizability of RCT findings to routine health care is less limited than frequently supposed. (*Psychiatric Services* 65:897–904, 2014; doi: 10.1176/appi.ps.201300309)

Compared with nonchronic forms of depression, chronic depression is associated with more severe functional impairment (1), greater use of health care, and higher hospitalization rates (2,3). Randomized con-

trolled trials (RCTs) have reported promising results for the efficacy of several pharmacological, psychotherapeutic, and combined pharmacological and psychotherapeutic interventions for chronic depression (4–11).

RCTs often focus on high internal validity—that is, whether potential group differences regarding the outcome can be attributed with a high degree of certainty to the received treatment. Yet the relevance of these trials for routine care depends on external validity (or generalizability)—that is, whether the results can be reasonably applied to a definable group of patients in a particular clinical setting in routine practice (12). External validity is often neglected in comparison with internal validity for several reasons. For example, drug licensing bodies, such as the U.S. Food and Drug Administration, do not require evidence that a drug has a clinically useful treatment effect or that a trial population is representative of routine clinical practice (12). In addition, guidance from ethics committees or funding agencies and reporting guidelines for RCTs emphasize internal validity over external validity (12). The focus on internal validity is often justified on the grounds that without internal validity, external validity or generalizability would be irrelevant or misleading (13). In recent years, practical clinical trials (PCTs) that aim to provide generalizable answers to important clinical questions without bias have been proposed; however, support for PCTs is still limited (14).

One strategy often used in RCTs to increase internal validity is to conduct research in a homogeneous sample of

---

Ms. von Wolff, Ms. Westphal, Dr. Härter, and Dr. Kriston are with the Department of Medical Psychology, University Medical Center Hamburg-Eppendorf, Hamburg, Germany (e-mail: a.von-wolff@uke.de). Mr. Jansen is with the Institute for Educational Quality Improvement, Humboldt University Berlin, Germany. Dr. Hölzel is with the Department of Psychiatry and Psychotherapy, University Medical Center Freiburg, Freiburg, Germany.

**Table 1**

Studies that examined differences between patients with depression who were eligible for a randomized controlled trial (RCT patients) and those who did not meet eligibility criteria (non-RCT patients)<sup>a</sup>

| Study                         | Sampling and treatment options                                                                                                                                                                                                                                                                                                                                                                                                                                           | Sample N | RCT patients (%)   | Eligibility criteria for RCT                                                                                                                                                                                                                                                                                                                                                                        | Main results                                                                                                                                                                                                                                                                                                                                                                                           |
|-------------------------------|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|----------|--------------------|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| Schindler et al., 2011 (22)   | Baseline interview and standard sessions (up to 5) of pretreatment phase for outpatient psychotherapy in Germany (N=1,584); primary diagnosis of depression (N=1,067); $\geq 6$ sessions of psychotherapy; BDI score indicating a depressive disorder; no incomplete data; treatments: cognitive-behavioral psychotherapy and cognitive psychotherapy                                                                                                                    | 338      | 76                 | No alcohol use or substance use disorders; no psychotic disorders, organic brain syndrome, developmental disorder, or dysthymic disorder; no suicide risk                                                                                                                                                                                                                                           | No baseline differences between RCT and non-RCT patients in age, gender, family status, years of education, psychotropic medication, previous treatment, number of treatment sessions, duration of therapy, somatoform disorders, eating disorders, or axis II comorbidity; more anxiety among RCT patients; both groups showed smaller treatment effect sizes than those reported in clinical trials. |
| van der Lem et al., 2011 (19) | Current major depressive disorder (N=1,653); at least one follow-up assessment of treatment outcome; stepped-care treatment options, including pharmacotherapy and psychotherapy                                                                                                                                                                                                                                                                                         | 626      | 17–25 <sup>b</sup> | No history of manic or hypomanic episodes, psychotic features, dysthymic disorder, anxiety disorders, eating disorders, attention-deficit hyperactivity syndrome, somatoform disorders, or borderline personality disorder; no high suicide risk; no alcohol or drug abuse or dependence (previous 6 months); HAM-D score $\geq 18$ ; duration of current episode $\geq 4$ weeks and $\leq 2$ years | No group differences between response and remission rates; no correlation between eligibility criteria and rates of response or remission                                                                                                                                                                                                                                                              |
| Wisniewski et al., 2009 (20)  | Age range of 18–75 years; single episode or recurrent major depressive disorder; HAM-D score $\geq 14$ ; no treatment resistance; not pregnant or breastfeeding; no bipolar disorder, psychosis, obsessive-compulsive disorder, eating disorder, or substance abuse requiring treatment; no medication other than study medication; no somatic comorbidity to contraindicate study medication (for example, seizures); N=4,177 consented to study; treatment: citalopram | 2,876    | 22                 | HAM-D score $\geq 19$ ; no more than one medical condition; no obsessive-compulsive disorder; no more than one concurrent psychiatric disorder; duration of current episode $\leq 24$ months                                                                                                                                                                                                        | For RCT patients: shorter duration of current episode, fewer with a family history of substance abuse, fewer with previous suicide attempts, and fewer with anxious and atypical symptoms. RCT patients had higher rates of response and remission (after analysis controlled for baseline differences).                                                                                               |
| Zetin and Hoepner, 2007 (21)  | Outpatients in private practice (N=817); age range 16–65 years; primary diagnosis of major depressive disorder or bipolar disorder; HAM-D score $\geq 7$ ; treatment: pharmacotherapy                                                                                                                                                                                                                                                                                    | 348      | 9                  | No bipolar or psychotic features; no suicidal ideation, comorbid axis I disorders, borderline personality disorder, or dysthymic disorder; HAM-D score $\geq 20$ ; duration of current episode $\geq 4$ weeks and $\leq 2$ years                                                                                                                                                                    | No between-group differences reported                                                                                                                                                                                                                                                                                                                                                                  |

*Continues on next page*

**Table 1***Continued from previous page*

| Study                       | Sampling and treatment options                                                                                           | Sample N | RCT patients (%) | Eligibility criteria for RCT                                                                                                                                                                                      | Main results                                                                                                                                                                                                                                 |
|-----------------------------|--------------------------------------------------------------------------------------------------------------------------|----------|------------------|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| Zimmerman et al., 2005 (18) | ≥18 years old; no bipolar or psychotic features; major depressive disorder; outpatient treatment (not further specified) | 599      | 21               | HAM-D score ≥20; no suicide risk; no recent diagnosis of substance abuse or dependence; no comorbid posttraumatic stress disorder, panic disorder, generalized anxiety disorder, or obsessive-compulsive disorder | Patients excluded from the RCT because of suicide risk or comorbidities had greater social impairment, more frequently missed work, longer episode duration, more previous suicide attempts, and more cluster B and C personality disorders. |

<sup>a</sup> Abbreviations: BDI, Beck Depression Inventory; HAM-D, Hamilton Rating Scale for Depression<sup>b</sup> Depending on assessment of personality

patients defined by strict eligibility criteria (15). However, the use of rigorous eligibility criteria to select patients for efficacy trials may limit the generalizability of results (16).

In recent years, the generalizability of findings from efficacy trials of interventions to treat depression has been questioned (17). Central to the question of generalizability is whether there are differences in clinical, demographic, and psychosocial characteristics and in treatment outcome between patients who would qualify for an antidepressant efficacy trial and patients who would not (18).

Between 9% and 76% of depressed outpatients (mainly those without chronic depression) who seek routine treatment would qualify for efficacy trials (18–22). Compared with these patients, those who would be excluded were found to be a more chronically ill group, with more previous episodes of depression, greater psychosocial impairment, more personality pathology (18), longer average duration of illness, higher rates of family history of substance abuse, more prior suicide attempts, and a greater number of anxious and atypical symptom features (20).

Inconsistent findings have been reported regarding differences in treatment outcome between patients who would or would not be included in a trial. More favorable outcomes were reported for patients who would typically be included in a trial of citalopram, compared with those who would typically be excluded (20). However,

another study did not report different between-group outcomes for psychotherapy (22). Treatment effects among patients in routine practice were found to be smaller than effect sizes reported in relevant RCTs (22). However, the influence of eligibility on treatment outcome was found to be rather small (19,23). Table 1 presents a detailed description of these studies, addressing differences between potentially included and excluded patients.

None of the studies mentioned above focused on chronically depressed inpatients. Therefore, little is known about the generalizability of findings from antidepressant efficacy trials for this subgroup in this setting. Compared with nonchronic depression, chronic depression is more often characterized by a longer duration of the depressive episode, a family history of mood disorders, and psychiatric comorbidity (24), and these characteristics are more often found among patients who are excluded from antidepressant efficacy trials (18,20), which is of particular relevance.

The objectives of this study were to identify typical eligibility criteria used in RCTs of treatments for chronic depression, to examine whether these criteria are met by patients with chronic depression treated in routine inpatient care, to examine whether patients meeting RCT eligibility criteria are treated differently in routine care, and to determine whether these patients respond differently to routine care than patients who do not meet RCT eligibility criteria.

## Methods

Eligibility criteria were extracted from RCTs of treatments for chronic depression. Trials were selected for this study if they included patients with chronic depression who were treated in at least one study arm with a combined psychotherapeutic and pharmacological intervention. We focused on combined interventions because these are recommended by current treatment guidelines for chronic depression (25–27) and because many patients in routine care are treated with more than one type of intervention (28). A total of eight RCTs were identified through a systematic database search (conducted in January 2010) including MEDLINE and EMBASE (4–11). A detailed description of the systematic search and of all included trials can be found elsewhere (29–31).

We conducted an exploratory secondary analysis of data from a large multicentre trial of routine inpatient depression treatment in Germany (32). The studied population included adult patients with any depressive disorder who received a minimum of three days of routine inpatient depression treatment in one of ten cooperating hospitals during the recruitment phase between December 2001 and February 2003. Because our secondary analysis focused only on chronic depression, we included patients with an onset of depression at least two years before admission to inpatient treatment who were currently treated either because of a “worsening of a chronic state” or “persistence of a long-lasting state”

**Table 2**

Eligibility criteria used in eight randomized controlled trials and patients in a routine care sample (N=231) who met or did not meet them

| RCT eligibility criteria                                                                                                                   | N of studies using criterion | Operationalization through routine data                                                                                                                                                    | Routine care sample  |     |                          |    |
|--------------------------------------------------------------------------------------------------------------------------------------------|------------------------------|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|----------------------|-----|--------------------------|----|
|                                                                                                                                            |                              |                                                                                                                                                                                            | Would meet criterion |     | Would not meet criterion |    |
|                                                                                                                                            |                              |                                                                                                                                                                                            | N                    | %   | N                        | %  |
| Ability to consent to study                                                                                                                | 7                            | Treated at own will                                                                                                                                                                        | 227                  | 98  | 4                        | 2  |
| Sufficient language skills                                                                                                                 | 2                            | German as first language or good German language skills                                                                                                                                    | 223                  | 97  | 8                        | 3  |
| Age between 18 and 75                                                                                                                      | 7                            | Age range 18–75                                                                                                                                                                            | 221                  | 96  | 10                       | 4  |
| No concurrent psychiatric disorders (most frequently schizophrenia, other psychotic disorder, bipolar disorder, or substance use disorder) |                              | Any ICD–10 diagnosis code indicating a mental or behavioral disorder due to psychoactive substance use; schizophrenia, schizotypal, or delusional disorders; or bipolar affective disorder |                      |     |                          |    |
|                                                                                                                                            | 8                            |                                                                                                                                                                                            | 194                  | 84  | 37                       | 16 |
| No concurrent somatic diseases                                                                                                             | 7                            | Somatic comorbidity                                                                                                                                                                        | 143                  | 62  | 88                       | 38 |
| Not pregnant or breastfeeding                                                                                                              | 5                            | Pregnant or lactating                                                                                                                                                                      | 231                  | 100 | 0                        | —  |
| No suicide risk                                                                                                                            | 4                            | High suicidal tendency <sup>a</sup> at admission or admission due to self-threatening behavior or attempted suicide just before admission                                                  | 163                  | 71  | 66                       | 29 |
| Minimum severity of depression at baseline                                                                                                 | 5                            | Hamilton Rating Scale for Depression score $\geq 16$ at admission                                                                                                                          | 170                  | 83  | 34                       | 17 |
| Specific diagnosis of chronic depression (for example, dysthymia)                                                                          | 8                            | na <sup>b</sup>                                                                                                                                                                            |                      |     |                          |    |
| Not previously treated with an intervention tested in the study                                                                            | 2                            | na <sup>b</sup>                                                                                                                                                                            |                      |     |                          |    |
| Nonresponse to an adequate intervention                                                                                                    | 3                            | na <sup>b</sup>                                                                                                                                                                            |                      |     |                          |    |
| Not currently treated with an intervention other than the study intervention                                                               | 5                            | na <sup>b</sup>                                                                                                                                                                            |                      |     |                          |    |
| No contraindications to the study intervention                                                                                             | 6                            | na <sup>b</sup>                                                                                                                                                                            |                      |     |                          |    |
| All above criteria                                                                                                                         |                              |                                                                                                                                                                                            | 58                   | 25  | 173                      | 75 |

<sup>a</sup> Assessed in accordance with the AMDP (Working Group for Methodology and Documentation in Psychiatry) system by the physician at admission

<sup>b</sup> Not applicable; no corresponding data collected in routine care

according to the documentation of the psychiatrist in charge at admission.

The psychiatrists documented the patients' demographic characteristics, history of depression, and psychopathology with the Psychiatric Basic Documentation System (33) and rated the patients' level of depression at admission and discharge with the Hamilton Rating Scale for Depression (HAM-D) (34). We operationalized each of the identified RCT eligibility criteria with data from the Psychiatric Basic Documentation System. On the basis

of these operationalizations, we calculated absolute and relative frequencies of whether the inpatients in routine care met the criteria. We also calculated the number and proportion of patients who met all eligibility criteria and who would thus be eligible to participate in RCTs (referred to below as "RCT patients") and the number and proportion who would be excluded from RCTs for any reason ("non-RCT patients").

We used three definitions of treatment outcome: response, defined as

symptom reduction between admission and discharge of at least 50%; remission, defined as severity of depression at discharge  $\leq 7$  points on the HAM-D; and absolute symptom reduction on the HAM-D between admission and discharge. Chi square tests and t tests were used to examine differences in treatment outcome between RCT patients and non-RCT patients.

We also compared the RCT and non-RCT groups on the effectiveness of various treatments strategies used

during the inpatient stay. More specifically, we focused on selective serotonin reuptake inhibitors (SSRIs), tricyclic antidepressants (TCAs), monoamine oxidase inhibitors (MAOIs), modern antidepressants (venlafaxine, mirtazapine, and reboxetine), and psychotherapeutic interventions. In a multigroup logistic regression model, we included the various treatment strategies as predictors of treatment outcome. We included all strategies in the same model to estimate the efficacy of a certain treatment strategy while statistically controlling for the effects of all other treatment strategies.

We decided not to control for any patient characteristics (such as baseline difference in depression severity) because we assumed that relevant patient characteristics influenced both the chance to be eligible for an RCT and the choice of treatment strategy in routine care. These baseline patient characteristics are therefore not considered to be “confounders” that need to be controlled to estimate unbiased treatment effects in our analysis. Rather, differences in these characteristics are inherent to the groups we want to compare (and if the differences lead to patients being eligible or not eligible, they also define the groups of interest). Our central research question was whether these differences themselves lead to differential effects of specific treatment strategies.

Regression coefficients were estimated separately in both groups (RCT and non-RCT patients) within the multigroup model, and group differences in regression coefficients were tested for statistical significance with z tests. Response rates were used as the outcome criterion, and odds ratios (ORs) were used as a measure of effect size. To compare ORs of response rates, multigroup logistic regression models were estimated. The differences in regression coefficients were tested for significance by z tests.

Analyses were performed using PASW Statistics for Windows, version 18.0, and Mplus 6.1 (35).

## Results

We identified 13 eligibility criteria used in eight RCTs of combined psychotherapeutic and pharmacological

**Table 3**

Characteristics of patients in a routine care sample (N=231) who met or did not meet all eligibility criteria for inclusion in randomized controlled trials<sup>a</sup>

| Characteristic                         | Met criteria<br>(N=58) |    | Did not meet criteria<br>(N=173) |    | p     |
|----------------------------------------|------------------------|----|----------------------------------|----|-------|
|                                        | N                      | %  | N                                | %  |       |
| Female                                 | 37                     | 64 | 107                              | 62 | .876  |
| Age (M±SD)                             | 50.1±11.2              |    | 52.1±13.4                        |    | .306  |
| Marital status                         |                        |    |                                  |    | .402  |
| Single                                 | 5                      | 9  | 28                               | 16 |       |
| Married                                | 39                     | 67 | 97                               | 56 |       |
| Divorced or separated                  | 9                      | 16 | 29                               | 17 |       |
| Widowed                                | 5                      | 9  | 19                               | 11 |       |
| Education                              |                        |    |                                  |    | .786  |
| Low (9 years)                          | 24                     | 45 | 84                               | 49 |       |
| Middle (10 years)                      | 15                     | 28 | 44                               | 26 |       |
| High (≥12 years)                       | 9                      | 17 | 32                               | 19 |       |
| None or other (<9 years)               | 5                      | 9  | 10                               | 6  |       |
| HAM-D score (M±SD) <sup>b</sup>        | 24.7±6.7               |    | 22.8±8.6                         |    | .102  |
| Duration of illness (M±SD years)       | 10.7±9.6               |    | 12.0±11.7                        |    | .439  |
| High risk of suicide                   | 0                      | —  | 66                               | 39 | <.001 |
| Psychiatric comorbidity <sup>c</sup>   | 0                      | —  | 37                               | 21 | <.001 |
| Somatic comorbidities                  | 0                      | —  | 88                               | 51 | <.001 |
| Treatment characteristic               |                        |    |                                  |    |       |
| Selective serotonin reuptake inhibitor | 22                     | 38 | 41                               | 24 | .042  |
| Tricyclic antidepressant               | 14                     | 24 | 53                               | 31 | .405  |
| Modern antidepressant <sup>d</sup>     | 24                     | 41 | 75                               | 43 | .878  |
| Monoamine oxidase inhibitors           | 4                      | 7  | 12                               | 7  | .999  |
| Psychotherapy                          | 37                     | 64 | 110                              | 64 | .999  |
| Combined treatment                     | 45                     | 78 | 121                              | 70 | .313  |
| Length of inpatient stay (M±SD days)   | 55.7±30.4              |    | 55.9±40.7                        |    | .971  |
| Outcome                                |                        |    |                                  |    |       |
| Response                               | 43                     | 77 | 94                               | 65 | .129  |
| Remission                              | 26                     | 46 | 73                               | 50 | .753  |
| Response and remission                 | 26                     | 46 | 71                               | 49 | .756  |
| Symptom reduction (M±SD)               | 15.8±7.9               |    | 13.5±9.1                         |    | .103  |

<sup>a</sup> Means were compared by t tests, and proportions were compared by chi square tests (Fisher's exact test).

<sup>b</sup> Hamilton Rating Scale for Depression. Possible scores range from 0 to 66, with higher scores indicating greater depression severity.

<sup>c</sup> Schizophrenia, other psychotic disorder, bipolar disorder, or substance use disorder

<sup>d</sup> Venlafaxine, mirtazapine, and reboxetine

interventions. All criteria were used in more than one study, and all studies used more than one eligibility criterion. Most studies explicitly mentioned the ability to provide written informed consent, a certain age range, and a specific diagnosis of chronic depression. Most studies excluded patients older than 75 and those who had various concurrent psychiatric or somatic disorders. The criteria are listed in Table 2, along with our operationalization of these criteria.

A total of 2,133 patients were included in the primary multicentre trial of routine inpatient depression

treatment in Germany (32), of which 231 were treated because of chronic depression and could be included in this secondary analysis. The mean±SD age of the 231 patients was 51.6±12.9 years, and more than half were female (62%). The mean inpatient stay for these patients was 55.8±38.3 days, and patients had a high severity of depression at admission (HAM-D score of 23.3±8.2). Most patients received a combination of psychotherapy and pharmacotherapy; seven (3%) were treated exclusively with psychotherapeutic interventions and 27 (12%) with pure pharmacological interventions.



**Table 4**

Treatment response of patients in routine care (N=231) who met all eligibility criteria for inclusion in randomized controlled trials (RCT patients) and those who did not (non-RCT patients)<sup>a</sup>

| Strategy                               | RCT patients |           | Non-RCT patients |          | p    |
|----------------------------------------|--------------|-----------|------------------|----------|------|
|                                        | OR           | 95% CI    | OR               | 95% CI   |      |
| Selective serotonin reuptake inhibitor | .51          | .12–2.44  | .99              | .36–2.69 | .483 |
| Tricyclic antidepressant               | 3.56         | .37–34.63 | 1.20             | .48–3.00 | .385 |
| Modern antidepressant <sup>b</sup>     | .67          | .14–3.13  | 1.58             | .61–4.07 | .484 |
| Monoamine oxidase inhibitors           | .23          | .02–2.60  | .62              | .15–2.60 | .353 |
| Psychotherapy                          | 1.00         | .24–4.17  | .56              | .27–1.16 | .471 |

<sup>a</sup> Reference group, psychotherapy. To compare odds ratios (ORs) of response rates, multigroup logistic regression models were estimated. The differences in regression coefficients were tested for significance (difference from zero) by z tests.

<sup>b</sup> Venlafaxine, mirtazapine, and reboxetine

The most frequent psychotherapeutic intervention was cognitive-behavioral therapy, followed by psychodynamic psychotherapy and interpersonal psychotherapy.

Of the 231 inpatients with chronic depression, only 25% met all 13 RCT eligibility criteria (Table 2). Four of the 13 criteria were met by nearly all patients: ability to give informed consent (98%), operationalized as being treated of their own will; sufficient language skills (97%); younger than 75 years (96%); and not pregnant at admission (100%). However, a rather high percentage of these patients would not be eligible for RCTs because of somatic (38%) or psychiatric (16%) comorbidities, a high risk of suicide (29%), or a rather low severity of depression (HAM-D  $\leq 16$ ) at admission (17%).

We did not detect any statistically significant differences at admission in demographic characteristics between RCT patients and non-RCT patients (Tables 3 and 4). Differences between RCT and non-RCT patients were found in clinical characteristics that were used as eligibility criteria. The non-RCT patients had various concurrent somatic or psychiatric disorders, which the RCT patients did not have, and a higher risk of suicide than the RCT patients. No statistically significant differences were found for baseline depression severity.

Of the RCT patients, 38% received SSRIs during their inpatient stay, compared with 24% of the non-RCT patients ( $p=.042$ ). No statistically significant

differences between the two groups were found in regard to treatment with TCAs, modern antidepressants, MAOIs, psychotherapy, combined psychotherapeutic and pharmacological interventions, and length of inpatient stay (Table 3).

At the end of inpatient treatment, 46% of the RCT patients and 50% of the non-RCT patients reached remission, which was not a significant difference. At this time, 77% of the RCT patients and 65% of the non-RCT patients were classified as responders, and the average reduction in HAM-D scores was  $15.8 \pm 7.9$  among RCT patients and  $13.5 \pm 9.1$  among non-RCT patients. Differences in treatment outcome at the end of inpatient treatment between RCT patients and non-RCT patients did not reach statistical significance (Table 3).

For the RCT and non-RCT groups, the ORs of response rates of patients who were or were not treated with a certain treatment (SSRIs, TCAs, modern antidepressants, MAOIs, or psychotherapy) were estimated. A comparison of the ORs indicated no statistically significant between-group differences in effectiveness (Table 4).

## Discussion

We identified a broad range of eligibility criteria typically used in RCTs that examine the effectiveness of combined psychotherapeutic and pharmacological interventions for chronic depression. Applying these eligibility criteria to a sample of inpatients with chronic depression in routine inpatient

care, we found that a rather low percentage of patients would be eligible for RCTs (25%). This finding is consistent with previous research involving patients with nonchronic forms of depression who were screened for eligibility for clinical trials or screened in outpatient psychiatric care (18–21). Patients would have been excluded mainly because of suicide risk, low severity of depression at admission, and concurrent psychiatric or somatic illnesses.

We compared patients who met RCT eligibility criteria and those who did not, and we did not find statistically significant differences in demographic characteristics, length of inpatient stay, treatment outcome, and effectiveness of certain antidepressants, except that the percentage of patients receiving SSRIs was slightly larger in the RCT group. Our results indicated substantial heterogeneity of the effectiveness of single treatment strategies both between groups (effect sizes differed strongly between groups) and within groups (large confidence intervals of ORs in each group). However, the group differences were not statistically significant. One explanation may be a lack of power resulting from the limited sample size. Another explanation may be that the effectiveness of single treatment strategies in inpatient depression treatment varies mainly within each group of patients rather than systematically between RCT and non-RCT groups. This explanation would indicate that factors other than eligibility criteria modify treatment effects in routine care.

It should be noted that we could examine the effectiveness of only five specific interventions. Additional treatment modalities used in inpatient treatment, such as occupational therapy or somatic interventions, were not examined. Interaction effects between the examined interventions could not be accounted for because of the small sample. It is therefore possible that response variability to specific interventions was “overridden” by exposure to intensive treatment in an inpatient setting.

Because this study was based on routine data from inpatient depression treatment, we could not apply all eligibility criteria typically used in RCTs, such as contraindications in regard to a

specific study drug or the formal assessment of a specific subgroup diagnosis. Thus, on the one hand, our operationalization of eligibility criteria was not as strict as those actually applied in RCTs. On the other hand, because we combined eligibility criteria extracted from multiple clinical trials, our criteria were stricter than those of a single study.

A major strength of this study was the use of routine data, which allowed investigation of the role of eligibility criteria in a sample of inpatients with chronic depression who did not undergo any kind of preselection, contrary to previous research, in which patients who sought treatment in clinical trials were included (20).

## Conclusions

Our study provides empirical evidence to inform the discussion of whether findings from clinical trials of treatments for chronic depression can be generalized to routine inpatient care. First, we showed that most of the inpatients with chronic depression in routine care would normally not be included in clinical trials. Second, we did not find evidence that treatment outcome significantly differed between patients who would and would not have been eligible for RCTs.

Our findings imply that even though most inpatients in routine care would not be suitable for clinical trials, the generalizability of RCT findings to routine care is less limited than frequently suspected. Furthermore, in clinical trials that apply commonly used eligibility criteria, a large number of patients (about four times the planned sample size) would need to be screened to yield a sufficient number of patients with chronic depression.

In our study we focused on possible effects of patient characteristics on the generalizability of findings from RCTs to routine care, which is a central aspect in assessing whether evidence from RCTs is useful in establishing evidence-based practice (13). However, for a broader evaluation of the generalizability of findings of clinical trials, further research is needed on other factors relevant to the evaluation of generalizability. Such research would examine the implementation of the intervention and maintenance of

effects—for example, by comparing whether antidepressants are delivered in the same way in clinical trials and routine care (dosage, duration, experience of staff members, and so forth). Another approach to account for the generalizability of findings from clinical trials is to choose a priori an appropriate research methodology for conducting clinical trials in fields as complex as the treatment of chronic depression in routine care, as proposed by the Medical Research Council guideline for developing and evaluating complex interventions (36).

## Acknowledgments and disclosures

For this study, trial-level data were collected in a project funded by a grant from the German Ministry of Education and Research (METACHRON; 01KG0923), and patient-level data were collected in a project supported by the German Research Network on Depression.

Dr. Hölzel has received a speaker's honorarium from AstraZeneca. The other authors report no competing interests.

## References

1. Satyanarayana S, Enns MW, Cox BJ, et al: Prevalence and correlates of chronic depression in the Canadian community health survey: mental health and well-being. *Canadian Journal of Psychiatry* 54:389–398, 2009
2. Berndt ER, Koran LM, Finkelstein SN, et al: Lost human capital from early-onset chronic depression. *American Journal of Psychiatry* 157:940–947, 2000
3. Gilmer WS, Trivedi MH, Rush AJ, et al: Factors associated with chronic depressive episodes: a preliminary report from the STAR-D project. *Acta Psychiatrica Scandinavica* 112:425–433, 2005
4. Browne G, Steiner M, Roberts J, et al: Sertraline and/or interpersonal psychotherapy for patients with dysthymic disorder in primary care: 6-month comparison with longitudinal 2-year follow-up of effectiveness and costs. *Journal of Affective Disorders* 68:317–330, 2002
5. de Mello MF, Myczcowski LM, Menezes PR: A randomized controlled trial comparing moclobemide and moclobemide plus interpersonal psychotherapy in the treatment of dysthymic disorder. *Journal of Psychotherapy Practice and Research* 10: 117–123, 2001
6. Keller MB, McCullough JP, Klein DN, et al: A comparison of nefazodone, the cognitive behavioral-analysis system of psychotherapy, and their combination for the treatment of chronic depression. *New England Journal of Medicine* 342:1462–1470, 2000
7. Kocsis JH, Gelenberg AJ, Rothbaum BO, et al: REVAMP Investigators: Cognitive behavioral analysis system of psychother-

apy and brief supportive psychotherapy for augmentation of antidepressant nonresponse in chronic depression: the REVAMP Trial. *Archives of General Psychiatry* 66:1178–1188, 2009

8. Markowitz JC, Kocsis JH, Bleiberg KL, et al: A comparative trial of psychotherapy and pharmacotherapy for “pure” dysthymic patients. *Journal of Affective Disorders* 89: 167–175, 2005
9. Miller IW, Norman WH, Keitner GI: Combined treatment for patients with double depression. *Psychotherapy and Psychosomatics* 68:180–185, 1999
10. Ravindran AV, Anisman H, Merali Z, et al: Treatment of primary dysthymia with group cognitive therapy and pharmacotherapy: clinical symptoms and functional impairments. *American Journal of Psychiatry* 156:1608–1617, 1999
11. Schramm E, Schneider D, Zobel I, et al: Efficacy of interpersonal psychotherapy plus pharmacotherapy in chronically depressed inpatients. *Journal of Affective Disorders* 109:65–73, 2008
12. Rothwell PM: External validity of randomised controlled trials: “to whom do the results of this trial apply?” *Lancet* 365:82–93, 2005
13. Green LW, Glasgow RE: Evaluating the relevance, generalization, and applicability of research: issues in external validation and translation methodology. *Evaluation and the Health Professions* 29:126–153, 2006
14. Tunis SR, Stryer DB, Clancy CM: Practical clinical trials: increasing the value of clinical research for decision making in clinical and health policy. *JAMA* 290:1624–1632, 2003
15. Stirman SW, DeRubeis RJ, Crits-Christoph P, et al: Are samples in randomized controlled trials of psychotherapy representative of community outpatients? A new methodology and initial findings. *Journal of Consulting and Clinical Psychology* 71: 963–972, 2003
16. Zimmerman M, Mattia JI, Posternak MA: Are subjects in pharmacological treatment trials of depression representative of patients in routine clinical practice? *American Journal of Psychiatry* 159:469–473, 2002
17. Parker G: Antidepressants on trial: how valid is the evidence? *British Journal of Psychiatry* 194:1–3, 2009
18. Zimmerman M, Chelminski I, Posternak MA: Generalizability of antidepressant efficacy trials: differences between depressed psychiatric outpatients who would or would not qualify for an efficacy trial. *American Journal of Psychiatry* 162:1370–1372, 2005
19. van der Lem R, van der Wee NJA, van Veen T, et al: The generalizability of antidepressant efficacy trials to routine psychiatric out-patient practice. *Psychological Medicine* 41:1353–1363, 2011
20. Wisniewski SR, Rush AJ, Nierenberg AA, et al: Can phase III trial results of antidepressant medications be generalized to clinical practice? A STAR\*D report. *American Journal of Psychiatry* 166:599–607, 2009

21. Zetin M, Hoepner CT: Relevance of exclusion criteria in antidepressant clinical trials: a replication study. *Journal of Clinical Psychopharmacology* 27:295–301, 2007
22. Schindler AC, Hiller W, Witthöft M: Benchmarking of cognitive-behavioral therapy for depression in efficacy and effectiveness studies: how do exclusion criteria affect treatment outcome? *Psychotherapy Research* 21:644–657, 2011
23. van der Lem R, de Wever WWH, van der Wee NJA, et al: 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* 12:192, 2012
24. Hölzel LP, Härter M, Reese C, et al: Risk factors for chronic depression: a systematic review. *Journal of Affective Disorders* 129: 1–13, 2011
25. Practice guideline for the treatment of patients with major depressive disorder. *American Journal of Psychiatry* 167(suppl 10):1–152, 2010
26. Depression: The Treatment and Management of Depression in Adults. National Clinical Practice Guideline 90. London, National Institute for Health and Clinical Excellence, 2009
27. Härter M, Klesse C, Bermejo I, et al: Evidence-based therapy of depression: S3 guidelines on unipolar depression [in German]. *Der Nervenarzt* 81:1049–1068, 2010
28. Fritze J: Prescriptions for psychotropic drugs: results and comments on the 2007 prescription drug report [in German]. *Der Nervenarzt* 79:1337–1345, 2008
29. Kriston L, von Wolff A, Hölzel LP: Effectiveness of psychotherapeutic, pharmacological, and combined treatments for chronic depression: a systematic review (METACHRON). *BMC Psychiatry* 10:95, 2010
30. von Wolff A, Hölzel LP, Westphal A, et al: Selective serotonin reuptake inhibitors and tricyclic antidepressants in the acute treatment of chronic depression and dysthymia: a systematic review and meta-analysis. *Journal of Affective Disorders* 144:7–15, 2013
31. von Wolff A, Hölzel LP, Westphal A, et al: Combination of pharmacotherapy and psychotherapy in the treatment of chronic depression: a systematic review and meta-analysis. *BMC Psychiatry*, 2012; doi 10.1186/1471-244X-12-61
32. Schneider F, Härter M, Brand S, et al: Adherence to guidelines for treatment of depression in in-patients. *British Journal of Psychiatry* 187:462–469, 2005
33. Cording C, Gaebel W, Spengler A, et al: The new psychiatric basic documentation: a recommendation of the DGPPN for quality insurance in inpatient treatment [in German]. *Spektrum der Psychiatrie und Nervenheilkunde* 24:3–41, 1995
34. Hamilton M: A rating scale for depression. *Journal of Neurology, Neurosurgery, and Psychiatry* 23:56–62, 1960
35. Muthén LK, Muthén BO: Mplus, Version 6.1. Los Angeles, Muthén and Muthén, 2011
36. Craig P, Dieppe P, Macintyre S, et al: Developing and evaluating complex interventions: the new Medical Research Council guidance. *BMJ*, 2008; doi 10.1136/bmj.a1655

## ***Psychiatric Services Invites Short Descriptions of Novel Programs***

*Psychiatric Services* invites contributions for Frontline Reports, a column featuring short descriptions of novel approaches to mental health problems or creative applications of established concepts in different settings.

Text should be 350 to 750 words. A maximum of three authors, including the contact person, can be listed; one author is preferred. References, tables, and figures are not used. Any statements about program effectiveness must be accompanied by supporting data within the text.

Material to be considered for Frontline Reports should be sent to one of the column editors: Francine Cournos, M.D., New York State Psychiatric Institute, 1051 Riverside Dr., Unit 112, New York, NY 10032 (e-mail: fc15@columbia.edu), or Stephen M. Goldfinger, M.D., Department of Psychiatry, SUNY Downstate Medical Center, Box 1203, 450 Clarkson Ave., Brooklyn, NY 11203 (e-mail: smgoldfingermd@aol.com).