Generalizability of Clinical Trial Results for Major Depression to Community Samples: Results From the National Epidemiologic Survey on Alcohol and Related Conditions

Carlos Blanco, M.D., Ph.D.; Mark Olfson, M.D., M.P.H.; Renee D. Goodwin, Ph.D., M.P.H.; Elizabeth Ogburn, M.S.; Michael R. Liebowitz, M.D.; Edward V. Nunes, M.D.; and Deborah S. Hasin, M.D.

Objective: Although emerging data indicate that sample composition may influence the effectiveness of mental health interventions, the extent to which subjects in clinical trials represent affected community samples remains unknown. The goal of this study was to assess the proportion of community-dwelling adults with major depressive episode (MDE) who would meet eligibility criteria for a traditional efficacy trial in patients with MDE.

Method: We applied a standard set of exclusion criteria used in clinical trials in patients with MDE to the 2001–2002 National Epidemiologic Survey for Alcohol and Related Conditions (NESARC), the largest psychiatric epidemiologic study in the United States to date (N = 43,093). Because individuals who seek treatment for a disorder may systematically differ from those who do not, we applied the criteria first to all individuals with a current diagnosis of MDE (N = 3119) (diagnosed according to DSM-IV) and then to the subsample of individuals who sought treatment (N = 1359).

Results: Among the full sample of individuals with MDE, 75.8% were excluded by one or more study eligibility criteria. Approximately two thirds (66.9%) of the subsample of those who sought treatment were excluded. The percentage of subjects excluded by individual study criteria ranged from 2.4% to 47.4% in the overall sample and 0% to 38.4% in the treatment-seeking sample. For both groups, the presence of comorbid, nondepressive, non–substance use Axis I disorders and the duration of the depressive episode excluded the largest percentage of individuals.

Conclusion: The design of traditional clinical trials tends to exclude a majority of individuals with MDE. Selection of exclusion criteria may have a powerful influence on the generalizability of study results. Clinical trials should explain the rationale for their exclusion criteria and estimate the impact of eligibility criteria on the generalizability of trial results.

(J Clin Psychiatry 2008;69:1276–1280)

Received Aug. 27, 2007; accepted Nov. 24, 2007. From New York State Psychiatric Institute and the Department of Psychiatry, College of Physicians and Surgeons of Columbia University, New York (Drs. Blanco, Olfson, Liebowitz, Nunes, and Hasin and Ms. Ogburn); and the Department of Epidemiology, Mailman School of Public Health, Columbia University, New York, N.Y. (Dr. Goodwin).

Supported by National Institutes of Health (NIH) grant P60 MD000206 to Dr. Olfson; NIH grants DA00482, DA015559, DA019606, and DA020783 to Dr. Blanco; and National Institute on Drug Abuse grants K24 DA022412-01 and K02 DA000288-10 to Dr. Nunes.

Financial disclosure is listed at the end of the article.

Corresponding author and reprints: Carlos Blanco, M.D., Ph.D., New York State Psychiatric Institute, 1051 Riverside Dr., Box 69, New York, NY 10032 (e-mail: cb255@columbia.edu).

Over the last several years, there has been an increased emphasis on practice of evidence-based medicine, generally understood as the application to clinical care of knowledge derived from clinical trials.[1–3] At the same time, emerging data indicate that sample composition may influence the results of clinical trials. This finding suggests that treatment effectiveness may be sensitive to specific inclusion and exclusion criteria.[4,5] As a consequence, there has been a call to quantify the generalizability of the results of clinical trials to the broader population of patients suffering from the disorder under study.[6–10]

Several approaches have been proposed to assess the representativeness of samples in clinical trials for a given disorder. One method is to compare the demographic and clinical characteristics and treatment response of trial participants recruited via advertisement with trial participants referred from clinical settings. With this approach, few group differences have been found, thus supporting the generalizability of efficacy trial results.[11–14] However, other methods have produced less reassuring results. For example, Stirman and colleagues[15] applied the exclusion criteria of clinical trials to patients seen in private psychotherapy practice. They found that nearly one half of the patients in the practices would have been excluded from research studies. In schizophrenia treatment trials, studies with placebo arms tend to enroll subjects with lower baseline severity than do trials with active comparator arms.[16]

A meta-analysis[17] found that exclusion rates for patients being considered for psychotherapy studies of depression, panic, and generalized anxiety disorder were 68%, 64%, and 65%, respectively. The exclusions were generally due to the presence of comorbid disorders that are typical of the disorder under study,[17] suggesting that the findings of those clinical trials may not be readily generalizable to patients under care. When Zimmerman and colleagues applied exclusion criteria common in pharmacotherapy clinical trials in patients with depression to a large sample of depressed patients attending a university outpatient psychiatric practice, only 14% of the outpatients met the eligibility criteria.[18] Furthermore, some of the exclusion criteria that excluded a higher percentage of patients, such as comorbid anxiety disorders, have been shown to influence the outcome for the treatment of major depressive episodes.[19–21] Therefore, the results of traditional clinical trials cannot be directly extrapolated to individuals in the community who suffer from depression.

The greater the proportion of individuals in need of treatment for the disorder who qualify for the trial, the more the study results will generalize to the affected population. The prior studies on the effects of exclusion criteria highlight the relevance of measuring the generalizability of clinical trial results and provide a framework for assessment. However, an important limitation of prior work has been reliance on local convenience samples to assess generalizability. From the standpoint of public policy, it is important to move beyond local, convenience, treatment-seeking samples to representative general population samples of individuals in need of treatment, providing a firmer basis from which the results can be more broadly generalized.

The present study assessed the effect of applying exclusion criteria commonly applied in clinical trials[22] to a large, nationally representative general population sample, the National Epidemiologic Survey for Alcohol and Related Conditions (NESARC). This approach provided a new method to estimate the population generalizability of clinical trial results. The NESARC is the largest epidemiologic study conducted to date in the United States. We examined the proportion of all cases of major depressive episode (MDE) in the NESARC that would be eligible if the exclusion criteria were applied to this sample. Because individuals who seek treatment for a disorder may systematically differ from those who do not,[23] we applied the criteria first to all individuals with a current diagnosis of MDE and then to the subsample of individuals who sought treatment.

METHOD

Source of Data

Data were drawn from the 2001–2002 NESARC, a survey of a nationally representative sample of the adult population of the United States (N = 43,093) conducted by the National Institute on Alcoholism and Alcohol Abuse (NIAAA) and described in detail elsewhere.[24–26] The target population was the civilian noninstitutionalized population, aged 18 years and older, residing in the United States. The overall survey response rate was 81%. Diagnoses were made according to the criteria of the DSM-IV using the NIAAA Alcohol Use Disorder and Associated Disabilities Interview Schedule-IV (AUDADIS-IV),[27] a fully structured diagnostic interview designed for experienced interviewers who are not clinicians. The reliability and validity of the AUDADIS-IV, including clinical reappraisal studies conducted by psychiatrists, are well documented in numerous national and international psychometric studies conducted in both clinical and general population studies.[28–30]

All potential NESARC respondents were informed in writing about the nature of the survey, the statistical uses of the survey data, the voluntary aspect of their participation, and the federal laws that rigorously provided for the strict confidentiality of the identifiable survey information. Those respondents consenting to participate after receiving this information were interviewed. The research protocol, including informed consent procedures, received full ethical review and approval from the U.S. Census Bureau and the U.S. Office of Management and Budget. This study was approved by the institutional review board of the New York State Psychiatric Institute.

Statistical Methods

Exclusion criteria commonly applied in clinical trials of treatments for MDE (see below in Clinical Trials Exclusion Criteria) were applied to individuals in the NESARC to determine the proportion of individuals from the general population with current MDE who would be eligible for the clinical trials. The same criteria were applied to the subset of individuals with current MDE who sought treatment to examine potential differences in eligibility between treatment-seeking and non–treatment-seeking individuals.

The NESARC weights each observation to correct for the complex sampling design, including differential probabilities of selection and nonresponse. We report percentages based on these weighted estimates. We used the SUDAAN[31] statistical software package to accommodate the sampling design and weights of the survey to calculate percentages and corresponding 95% confidence intervals.

Clinical Trials Exclusion Criteria

We used the exclusion criteria proposed by Zimmerman and colleagues[18] because they represent the best systematized set of criteria used in clinical trials in patients with MDE. These criteria are listed in Table 1. The percentages of individuals excluded by criteria 1 and 2 and 6 through 8 were estimated from data collected by the AUDADIS-

Table 1. Estimated Percentage of Adults With Major Depressive Episode in the NESARC Excluded by Traditional Efficacy Eligibility Criteria

Exclusion Variable	Full Sample (N = 3119), % (95% CI)	Treatment-Seeking Sample (N = 1359), % (95% CI)
Traditional efficacy eligibility criteria[a]		
1. History of mania or hypomania	17.4 (15.7 to 19.1)	0.0
2. Current psychotic features	2.4 (1.7 to 3.1)	1.4 (0.9 to 2.2)
3. Significant risk of suicide	8.9 (7.7 to 10.1)	6.6 (5.3 to 8.0)
4. Alcohol/drug use disorder in the last year	8.8 (7.6 to 10.0)	8.3 (6.9 to 9.9)
5. Score < 18 on HAM-D	…	…
6. Comorbid dysthymic disorder	16.0 (14.3 to 17.7)	14.4 (12.4 to 16.6)
7. Other past-year comorbid Axis I disorders[b]	47.4 (45.1 to 49.7)	34.1 (31.2 to 37.1)
8. Episode duration of < 4 weeks or > 2 years	40.3 (37.9 to 42.6)	38.4 (35.5 to 41.3)
9. Presence of borderline personality disorder	…	…
Excluded by any criterion	75.8 (73.9 to 77.8)	66.9 (64.1 to 69.6)

[a]Derived from Zimmerman et al.[18]
[b]Includes bipolar disorder, panic disorder, agoraphobia, social anxiety disorder, specific phobia, and generalized anxiety disorder.
Abbreviations: HAM-D = Hamilton Rating Scale for Depression, NESARC = National Epidemiologic Survey for Alcohol and Related Conditions.
Symbol: … = not in dataset.

IV. Criterion 3 was considered met if the person reported suicidal ideation in the last year, the time frame used by the AUDADIS-IV when assessing for the presence of "current" symptoms. Criterion 4 was applied using a 12-month rather than 6-month time frame. Information to approximate criteria 5 and 9 was not available in the NESARC.

Analysis Plan

We first determined the percentage (and 95% confidence interval) of survey respondents who would be excluded by individually applying each criterion of the clinical trials.[18] Because individuals might have been excluded by more than 1 criterion, we also calculated the overall percentage of subjects who would have been excluded by the simultaneous application of all of the measurable criteria.

RESULTS

The percentage of subjects excluded by at least 1 criterion was 75.8% in the full sample of 3119 individuals who met DSM-IV criteria for MDE and 66.9% in the subsample of 1359 individuals who sought treatment. The percentage of subjects excluded due to the application of a single criterion ranged from 2.4% to 47.4% in the overall sample of individuals with MDE and from 0% to 38.4% among those who sought treatment (Table 1). The proportion of individuals excluded by each criterion was similar in the full MDE sample and in the subsample of treatment-seeking individuals. For both respondent groups, the presence of comorbid, nondepressive, non–substance use Axis I disorders and the duration of the episode were the 2 criteria excluding the highest percentage of individuals. Presence of current psychotic features was the criterion excluding the lowest percentage of individuals in the overall sample, while history of mania or hypomania was the criterion least likely to exclude treatment-seeking individuals. More than two thirds of respondents from both the full MDE sample and the treatment-seeking subsample would have been excluded by one or more of the study criteria.

DISCUSSION

Consistent with earlier research, the results of this study suggest that traditional clinical trials tend to exclude, by design, a majority of individuals with MDE. In a previous study,[18] approximately 1 in 7 patients (14%) seeking treatment at a university mental health clinic was determined to be eligible for inclusion in a clinical trial. High rates of exclusion have also been reported from several,[32,33] although not all,[34] clinical trials of antidepressant efficacy. By employing a national epidemiologic sample as the reference group, we sought to place the implications of exclusion criteria selection within a broad public health context.

We found that more than two thirds of individuals with MDE would be excluded from traditional antidepressant clinical trials, regardless of whether the full sample or only those who sought treatment were included in the analysis. Our findings raise questions about the generalizability of clinical trial results to individuals with MDE in the community and have implications for the design of clinical trials. Selection of exclusion criteria may have a powerful influence on the generalizability of study results. Specification a priori of the goals of the study and estimation of the proportion of individuals ineligible for the trial via application of each exclusion criterion to the target population would assist study design. Such an approach may help researchers to weigh the trade-offs between statistical power (e.g., through use of highly homogeneous samples) and representativeness of the study sample, depending on whether the emphasis of the study is efficacy

or effectiveness.[10] Efficacy studies may benefit from relatively stringent inclusion/exclusion criteria in order to maximize detection of drug-placebo differences. By contrast, effectiveness studies place a larger emphasis on the generalizability of their findings by minimizing exclusion criteria.

We selected the NESARC for our study because it is the largest epidemiologic study with information on MDE in U.S. adults. However, the NESARC may not be the appropriate referent for all studies. In general, selection of the survey used as the reference sample should be guided by the population to which the clinical trial intends to generalize its results. For example, the forthcoming National Comorbidity Survey for Adolescents may be a better epidemiologic sample to assess the generalizability of clinical trial designs for adolescents. Similarly, studies focusing on special populations, e.g., Hispanics, may seek to limit the referent to only those individuals, e.g., using the Hispanic subsample of the NESARC rather than using the full NESARC sample.

Another important aspect in the evaluation of the generalizability of clinical trial results is whether a trial focuses on the eligibility of participants (i.e., "a priori" or prospective generalizability) or on the characteristics of the sample actually recruited ("a posteriori" generalizability). In this study, we focused on the exclusion criteria. This approach focuses on the prospective generalizability of clinical trial results but provides no information on the subjects who actually enter those studies. Our method establishes an upper boundary to the generalizability of clinical trial results. However, most studies fail to achieve representative participation by gender, race/ethnicity, and other sociodemographic and clinical variables, resulting in additional loss of representativeness of the study samples. As a result, the actual (a posteriori) generalizability of the studies is almost always lower than their a priori generalizability. We are currently developing and testing approaches to evaluate the a posteriori generalizability of clinical trial results. Studies that do not reach certain levels of a posteriori generalizability may have limited influence in shaping evidence-based practice.

The current study has several limitations. First, we adopted specific conventions to translate clinical criteria to the NESARC sample. Different conventions might have yielded different exclusion estimates. For example, we excluded all individuals with suicidal ideation in the last 12 months because that was the closest question and time frame available in the epidemiologic studies, but many clinical trials exclude only patients with *serious, recent* suicidal ideation. The percentage of individuals excluded by this criterion was relatively small compared to the percentage excluded by other criteria. Nevertheless, development of standardized procedures to operationalize clinical criteria in epidemiologic data might help improve future generalizability estimates.

Second, 2 of the exclusion criteria could not be operationalized. Thus, the number of subjects excluded from clinical trials is likely to be somewhat higher than estimated by our study. For example, Zimmerman and colleagues[35] found that the most common cutoff score for the Hamilton Rating Scale for Depression ([HAM-D] < 14) would have excluded approximately 32% to 47% of their sample.[35] A more precise operationalization in future epidemiologic samples of the exclusion criteria used in clinical trials may help prevent underestimating the proportion of individuals ineligible for those trials.

Third, we focused on the exclusion criteria applied to traditional efficacy trials rather than effectiveness trials because efficacy trials compose the vast majority of the clinical trials for major depressive disorder. We also attempted to assess the percentage of individuals with MDE who would be excluded by the Sequenced Treatment Alternatives to Relieve Depression (STAR*D) study, the largest effectiveness study of MDE treatment to date, which was designed to include a broadly representative sample of individuals with MDE.[36,37] However, more than half of the exclusion criteria used by STAR*D (e.g., prior intolerability to or lack of response to selective serotonin reuptake inhibitors, lack of response to ≥ 16 sessions of cognitive-behavioral therapy in the current episode, lack of response to ≥ 7 sessions of electroconvulsive therapy) could not be operationalized using the NESARC data. This finding underscores the need for a dialogue between clinicians, epidemiologists, and health services researchers on how to systematically collect data that may allow the comparison of subjects participating in clinical trials with those in the general population. Without that collaboration, evaluation of the generalizability of clinical trial results will remain an elusive goal for researchers and an ongoing concern for patients, clinicians, and policy-makers.

Fourth, we applied all exclusion criteria that could be operationalized with the NESARC, although each efficacy trial applied only a subset of these criteria. Thus, the percentage of individuals excluded by "any exclusion criteria" in our study may constitute an overestimate for most trials. However, a recent review of 39 efficacy studies in major depressive disorder[35] found that the frequency of use of the applied efficacy exclusion criteria ranged from 20.5% (presence of borderline personality disorder) to 92.3% (minimum score on the HAM-D).

Despite these limitations, our study suggests that application of traditional eligibility criteria of a clinical trial to an epidemiologic sample provides a means of quantifying the prospective generalizability of clinical trial results. The results confirm that traditional clinical efficacy trials in patients with depression are likely to involve highly selected samples. Future research should evaluate the applicability of this method to other types of studies and samples and assess the accuracy of alternative methods to measure the generalizability of clinical trial results.

Financial disclosure: **Dr. Olfson** has received grant support from Eli Lilly, Bristol-Myers Squibb, and AstraZeneca; is a consultant for Bristol-Myers Squibb, Eli Lilly, Pfizer, and McNeil; is a member of the advisory boards for Eli Lilly and Bristol-Myers Squibb; and is a member of the speakers bureau for Janssen. **Dr. Liebowitz** has equity ownership of ChiMatrix and the Liebowitz Social Anxiety Scale (LSAS); has received grant/research support from Pfizer, Eli Lilly, GlaxoSmithKline, Novartis, Wyeth, and Avera; is a consultant for Avera, AstraZeneca, Pfizer, Tikvah, Eli Lilly, Zars, and Wyeth; has received honoraria from Wyeth; is a member of the speakers/advisory boards for Jazz, AstraZeneca, and Zars; licenses LSAS software to GlaxoSmithKline, Pfizer, Avera, Tikvah, Eli Lilly, Indevus, Servier, and Wyeth; and has clinical trial contracts with Pfizer, GlaxoSmithKline, AstraZeneca, Forest, Tikvah, Avera, Eli Lilly, Novartis, Sepracor, Horizon, and Johnson & Johnson. **Dr. Nunes** is a member of the speakers/advisory boards for Alkermes/Cephalon and works as an investigator in a Titan study. **Drs. Blanco**, **Goodwin**, and **Hasin** and **Ms. Ogburn** report no additional financial or other relationships relevant to the subject of this article.

REFERENCES

1. Oxman AD, Sackett DL, Guyatt GH, for the Evidence-Based Medicine Working Group. Users' guides to the medical literature, 1: how to get started. JAMA 1993 Nov;270(17):2093–2095
2. Bucher HC, Guyatt GH, Cook DJ, et al, for the Evidence-Based Medicine Working Group. Users' guides to the medical literature, 19: applying clinical trial results, A: how to use an article measuring the effect of an intervention on surrogate end points. JAMA 1999 Aug;282(8):771–778
3. Geddes JR, Game D, Jenkins NE, et al. What proportion of primary psychiatric interventions are based on evidence from randomised controlled trials? Qual Health Care 1996;5:215–217
4. Khan A, Kolts RL, Thase ME, et al. Research design features and patient characteristics associated with the outcome of antidepressant clinical trials. Am J Psychiatry 2004;161:2045–2049
5. Khan A, Brodhead AE, Kolts RL, et al. Severity of depressive symptoms and response to antidepressants and placebo in antidepressant trials. J Psychiatr Res 2005;39:145–150
6. March JS, Silva SG, Compton S, et al. The case for practical clinical trials in psychiatry. Am J Psychiatry 2005;162:836–846
7. Geddes J, Carney S. Recent advances in evidence-based psychiatry. Can J Psychiatry 2001;46:403–406
8. Wells KB. Treatment research at the crossroads: the scientific interface of clinical trials and effectiveness research. Am J Psychiatry 1999;156:5–10
9. 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. Am J Psychiatry 2005;162:1370–1372
10. Van Spall HG, Toren A, Kiss A, et al. Eligibility criteria of randomized controlled trials published in high-impact general medical journals: a systematic sampling review. JAMA 2007;297:1233–1240
11. Rabinowitz J, Bromet EJ, Davidson M. Are patients enrolled in first episode psychosis drug trials representative of patients treated in routine clinical practice? Schizophr Res 2003;61:149–155
12. Brauzer B, Goldstein BJ. Symptomatic volunteers: another patient dimension for clinical trials. J Clin Pharmacol New Drugs 1973;13:89–98
13. Breckenridge JS, Zeiss AM, Breckenridge JN, et al. Solicitation of elderly depressives for treatment outcome research: a comparison of referral sources. J Consult Clin Psychol 1985;53:552–554
14. Rapaport MH, Frevert T, Babior S, et al. A comparison of demographic variables, symptom profiles, and measurements of functioning in symptomatic volunteers and an outpatient clinical population. Psychopharmacol Bull 1995;31:111–114
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. J Consult Clin Psychol 2003 Dec;71(6):963–972
16. Mohr P, Czobor P. Subject selection for the placebo- and comparator-controlled trials of neuroleptics in schizophrenia. J Clin Psychopharmacol 2000;20:240–245
17. 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. J Consult Clin Psychol 2001;69:875–899
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. Hegel MT, Unützer J, Tang L, et al. Impact of comorbid panic and posttraumatic stress disorder on outcomes of collaborative care for late-life depression in primary care. Am J Geriatr Psychiatry 2005;13:48–58
20. Frank E, Shear MK, Rucci P, et al. Influence of panic-agoraphobic spectrum symptoms on treatment response in patients with recurrent major depression. Am J Psychiatry 2000;157:1101–1107
21. Mittal D, Fortney JC, Pyne JM, et al. Impact of comorbid anxiety disorders on health-related quality of life among patients with major depressive disorder. Psychiatr Serv 2006;57:1731–1737
22. Zimmerman FJ. On the standard gamble. Arch Pediatr Adolesc Med 2003;157:225–227
23. Mojtabai R, Olfson M, Mechanic D. Perceived need and help-seeking in adults with mood, anxiety, or substance use disorders. Arch Gen Psychiatry 2002;59:77–84
24. Grant B, Moore TC, Shepard J, et al. Source and accuracy statement: Wave 1 National Epidemiologic Survey on Alcohol and Related Conditions (NESARC). Bethesda, Md: National Institute on Alcohol Abuse and Alcoholism; 2003. Available at: www.niaaa.nih.gov. Accessibility verified April 24, 2008.
25. Grant BF, Stinson FS, Dawson DA, 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
26. Hasin DS, Goodwin RD, Stinson FS, et al. Epidemiology of major depressive disorder: results from the National Epidemiologic Survey on Alcoholism and Related Conditions. Arch Gen Psychiatry 2005;62:1097–1106
27. Grant BF, Dawson DA, Stinson FS, et al. The Alcohol Use Disorder and Associated Disabilities Interview Schedule-IV (AUDADIS-IV): reliability of alcohol consumption, tobacco use, family history of depression and psychiatric diagnostic modules in a general population sample. Drug Alcohol Depend 2003;71:7–16
28. Grant BF, Harford TC, Dawson DA, et al. The Alcohol Use Disorder and Associated Disabilities Interview Schedule (AUDADIS): reliability of alcohol and drug modules in a general population sample. Drug Alcohol Depend 1995;39:37–44
29. Hasin D, Carpenter KM, McCloud S, et al. The Alcohol Use Disorder and Associated Disabilities Interview Schedule (AUDADIS): reliability of alcohol and drug modules in a clinical sample. Drug Alcohol Depend 1997;44:133–141
30. Canino G, Bravo M, Ramirez R, et al. The Spanish Alcohol Use Disorder and Associated Disabilities Interview Schedule (AUDADIS): reliability and concordance with clinical diagnoses in a Hispanic population. J Stud Alcohol 1999;60:790–799
31. Research Triangle Institute. Software for Survey Data Analysis (SUDAAN), version 9.0. Research Triangle Park, NC: Research Triangle Institute; 2004
32. 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 Sep;64(9):1091–1093
33. Posternak MA, Zimmerman M, Keitner GI, et al. A reevaluation of the exclusion criteria used in antidepressant efficacy trials. Am J Psychiatry 2002;159:191–200
34. Partonen T, Sihvo S, Lonnqvist JK. Patients excluded from an antidepressant efficacy trial. J Clin Psychiatry 1996 Dec;57(12):572–575
35. 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
36. Trivedi MH, Fava M, Wisniewski SR, et al. Medication augmentation after the failure of SSRIs for depression. N Engl J Med 2006;354:1243–1252
37. Trivedi MH, Rush AJ, Wisniewski SR, et al. Evaluation of outcomes with citalopram for depression using measurement-based care in STAR*D: implications for clinical practice. Am J Psychiatry 2006;163:28–40

