Accurate Meta-Analytical Assessment of “True Antidepressant Effects” Needed

Sir: In their recent article, Posternak and Zimmerman1 question whether there is a several-week delay in “true antidepressant effect.” They propose to address this issue by a meta-analysis to determine (1) “whether significant drug-placebo separation occurs during the first 2 weeks of treatment” and (2) “whether the timing of response to antidepressant medication and placebo is distinct.”

Since the issue centers on “true antidepressant effects,” measures and analyses must address this salient point. There is no argument that antidepressant drugs can produce immediate effects (e.g., sedation, weight gain, dry mouth), but these are not “true antidepressant effects.” Rather, they are adventitious toxicities that are not on the causal path to mood normalization.

The initial analyses focused on placebo-controlled trials lasting at least 4 weeks with at least b.i.w.e. (every other week) measures of the Hamilton Rating Scale for Depression (HAM-D). Various debatable exclusion criteria were used, but, for this discussion, these are beside the point. Their sample consisted of 47 studies. The discussion is obscured by a reliance on percentage reductions in HAM-D scores. The HAM-D is ordinal, and percentage change is not appropriate for such scales. Further, floor effects are ignored.

The simplest analysis would be to take the reported HAM-D scores at baseline and 2 weeks within each study and develop a contrast for drug versus placebo by any one of several standard methods (e.g., Wilcoxon signed rank test) and then aggregate these values by standard meta-analytic methods.

If there were no difference, that would end the discussion. If there were a difference, it would still be necessary to show that it was relevant to “true antidepressant effect.”

The authors’ actual analyses are difficult to understand. Establishing a mean baseline score and difference scores across all studies that account for sample size differences, as well as drug only, seems needlessly complex and incorporates many shaky assumptions about trial parallelism. Whatever is meant by “drug-placebo differences, after adjusting for the fewer number of subjects in weeks 5 and 6”1 when presenting drug-placebo differences for weeks 1 and 2 is entirely obscure.

The distinction between “true antidepressant” and other drug effects is only attempted by the authors’ limiting their analysis to what they consider nonsedating antidepressants. Peculiarly, they do not exclude paroxetine, so their judgment about nonsedative medications is arguable. In any case, this analysis does not speak to the manifold other irrelevant drug effects.

They ignore even the confounding soporific issue for their Clinical Global Impressions scale (CGI) analysis. Nor do they make any effort to see if weekly CGI responses represent transient fluctuations or maintained effects, as the Columbia group has.2–9

We conclude that, to address the problem of the onset of “true antidepressant effects,” such a meta-analysis is inappropriate.

Dr. Quitkin has been a consultant for Organon, has received honoraria from Pfizer, has been on the speakers or advisory board for Eli Lilly, and is a major stock shareholder in Cyberonics. Dr. Stewart has received grant/research support from Eli Lilly, Pfizer, Organon, and Glaxo; has received honoraria from Eli Lilly, Organon, Pfizer, Forest, Shire, and Somerset; and has been on the speakers or advisory boards for Organon and Shire. Dr. Klein has been a consultant for Sepracor and has been on the speakers or advisory boards for Vela Pharmaceuticals and VivoMetrics. Drs. McGrath and Ross report no financial or other relationship relevant to the subject of this letter.

REFERENCES

2. Quitkin FM, Rabkin JD, Ross D, et al. Duration of antidepressant drug treatment: what is an adequate trial? Arch Gen Psychiatry 1984;41:238–245

Frederic M. Quitkin, M.D.
Patrick J. McGrath, M.D.
Jonathan W. Stewart, M.D.
Donald F. Klein, M.D.
Donald C. Ross, Ph.D.
Department of Therapeutics
Columbia University College of Physicians and Surgeons
New York State Psychiatric Institute
New York, New York

Drs. Posternak and Zimmerman Reply

Sir: We thank Dr. Quitkin and colleagues from Columbia University for their comments regarding our article.1 As we point out in the article, the investigators from Columbia University were the pioneers in advancing our understanding of the timing of the antidepressant response. The findings from our meta-analysis, however, largely contradict their delayed antidepressant response theory, and Dr. Quitkin and colleagues reject our findings after arguing that the meta-analysis we conducted was “inappropriate.”

The claim of inappropriateness is supported by essentially 6 arguments, to each of which we offer a brief response:

1. Examining percentage change in scores on the Hamilton Rating Scale for Depression (HAM-D) is inappropriate because the HAM-D is ordinal. This comment is puzzling. Examining the percentage reduction in HAM-D scores is the standard for evaluating drug efficacy, and a 50% reduction in baseline HAM-D scores is the standard used in the field for defining antidepressant response. All 47 trials in our meta-analysis incorporated analyses based on percentage reductions in HAM-D scores.
2. Floor effects are ignored. Presumably, this comment refers to those subjects who dramatically improve and have little room for further improvement. The published data
LETTERS TO THE EDITOR

were not presented in a way that would have allowed us to look at floor effects. However, the central finding that significant drug-placebo differences were found in the first 2 weeks of treatment is not undermined in any way by the potential for floor effects, since the delayed antidepressant response theory would predict equivalent nonspecific improvement in both the drug and placebo cohorts.

3. The analyses are needlessly complex. We are somewhat embarrassed to admit that all analyses were conducted using a simple, hand-held calculator using the addition, subtraction, multiplication, and division keys. The only statistical tests conducted were c² tests and t tests. We would frankly find it more difficult to defend the charge that our analyses were overly simplistic.

4. The analyses incorporate many shaky assumptions about trial parallelism. Dr. Quitkin and colleagues do not indicate which assumptions they are referring to. However, there is an inherent risk when conducting any meta-analysis that combining results from studies with disparate methodologies and distinct study populations may lead to inappropriate conclusions. This risk holds true, of course, with all meta-analyses. In this case, however, the design of antidepressant efficacy trials has remained remarkably stagnant over the years, and trial investigators have been criticized for “slavishly” adhering to a monolithic research model. Thus, the trials included were remarkably homogeneous in their design, which we believe makes them particularly well suited for a meta-analysis such as this.

5. The analyses can not tease apart “true antidepressant effects” from “irrelevant drug effects.” The argument here pertains to the adventitious side effects that some antidepressants have, such as mirtzapine’s ameliorating insomnia and anorexia or bupropion’s ameliorating anergia. When a patient says he or she is eating better, is sleeping better, and has more energy, better concentration, and consequently a better mood, are these incidental “side effects” or the results of a true antidepressant effect? No method known to us can make this distinction, nor is it even clear to us that such a distinction is valid or worth making. This point notwithstanding, a subanalysis was performed that excluded the most sedating antidepressants, and this analysis confirmed that the early advantage of antidepressant therapy over placebo could not be attributed to the soporific effects of sedating antidepressants. (An a priori decision was made to include paroxetine because it is only mildly sedating. Our results would undoubtedly have been unchanged if paroxetine had been excluded.) A secondary analysis that relied only on the Clinical Global Impressions scale (CGI)—the same instrument used by the Columbia group—confirmed an early clinically significant antidepressant effect. If this instrument, too, is deemed unacceptable for establishing a true antidepressant effect, then we would ask the Columbia group whether any instrument is capable of discerning a true drug effect.

6. The subanalysis using the CGI instrument does not examine whether the effect is maintained. This type of longitudinal data was not presented by the trial investigators and therefore was not available to us. However, this information is largely irrelevant. We found that clinicians who were blind to treatment assignment were significantly more likely to rate subjects receiving an antidepressant medication as much improved during the first 1 or 2 weeks of treatment compared with subjects receiving placebo. This finding is inconsistent with the delayed antidepressant response theory, which predicts equivalent response rates during this time frame.

In sum, Dr. Quitkin and colleagues reject our conclusions after questioning the validity of the instruments and methodologies employed in our meta-analysis. However, our analyses followed the standard of instrumentation and methodology used in nearly all antidepressant trials, except that we pooled the results of dozens of trials. If this methodology is judged invalid, then the efficacy of antidepressants themselves would need to come into question as well. We conclude that if antidepressants work, they begin to work quickly, i.e., within the first 1 or 2 weeks of treatment. We are unable to reconcile our results with the findings from the Columbia group, but would suggest that further research is needed in order to definitively determine when antidepressants begin to work and why some patients take longer than others to respond.

Drs. Posternak and Zimmerman report no financial or other relationship relevant to the subject of this letter.

REFERENCES

2. Thase ME. How should efficacy be evaluated in randomized clinical trials of treatments for depression? J Clin Psychiatry 1999;60 (suppl 4):23–31

Michael A. Posternak, M.D.
Mark Zimmerman, M.D.
Department of Psychiatry and Human Behavior
Brown University School of Medicine
Providence, Rhode Island

Differential Modulation of Cerebrospinal Fluid Neurotrophins in Patients With Atopic Dermatitis Who Attempted Suicide

SIR: Controversy exists regarding the relationship that levels of neurotrophins in cerebrospinal fluid (CSF) or brain have with depression and/or suicide. Hock et al.1 state that neurotrophin-3 (NT-3) expression is elevated in elderly patients with depression, whereas Dwivedi et al.2 report that messenger RNA levels of brain-derived neurotrophic factor (BDNF) were found, in postmortem examinations, to be reduced in the brains of adults who committed suicide. On the other hand, Hadiconstantinou et al.3 state that plasma nerve growth factor (NGF) levels are elevated in depressed subjects. In contrast, Kimata4 states that kissing reduces plasma neurotrophin-4 (NT-4) levels. Therefore, I hypothesize that neurotrophins in CSF, in particular NGF, NT-3, NT-4, and BDNF, may be modulated by depression or suicide attempt. From 2001 through 2004, a total of 20 atopic dermatitis patients (12 female and 8 male, aged 13–16 years) attempted suicide and were immediately transferred to my hospital, which led to the current study of CSF levels of neurotrophins.

Method. Severity of atopic dermatitis was assessed by the SCORAD index.5 All 20 patients had moderate skin symptoms (mean [SEM] SCORAD index for sections A and B = 28 [2]), and they had been treated by local dermatologists with no subsequent improvement in skin condition. None of them took any antidepressant medication. After informed consent was obtained from the patients’ parents, CSF was obtained by lumbar puncture immediately after admission. Severity of depression was assessed by the Montgomery-Asberg Depression Rating Scale (MADRS).6 At admission, all 20 atopic dermatitis patients were
In addition to the subjects reported here, from 2004 through 2005 my colleagues and I treated 5 bronchial asthma patients who had attempted suicide (3 female and 2 male patients, aged 14–17 years). CSF levels (mean [SEM] pg/mL) of neurotrophins were as follows: NGF, 24.4 (3.1); NT-3, 34.5 (7.2); NT-4, < 9.4; and BDNF, 28.3 (3.6). Thus, change in CSF levels of neurotrophins may be due to depression, but not due to atopic dermatitis or bronchial asthma. However, the prevalence of suicidal ideation in bronchial asthma patients, mild atopic dermatitis patients, moderate atopic dermatitis patients, and severe atopic dermatitis patients is 0.17%, 0.21%, 6.00%, and 19.60%, respectively. Whether atopic dermatitis may induce depression is currently under investigation.

Collectively, these results indicate that CSF neurotrophins may be differentially involved in, and NT-4 and BDNF may be protective against, suicide attempts and/or depression and that measurement of CSF neurotrophins may be useful for the assessment of severity of depression or suicidal ideation.

Dr. Kimata reports no financial or other relationship relevant to the subject of this letter.

Hajime Kimata, M.D., Ph.D.
Department of Allergy
Satou Hospital, Hirakata-City
Osaka Prefecture, Japan

The ABCs of Suicide

SIR: One of psychiatry’s emergencies is the evaluation of a patient brought to the emergency room after a suicide attempt. To collect all relevant clinical information in this scenario, I have used a mnemonic that is modeled after the ABCs 

A—Airway, B—Breathing, C—Circulation

The alphabet is easily recalled even under the busiest of circumstances.

The ABCs of suicide is informed by the need for a patient-specific formulation using Mann and colleagues’ stress-diathesis model of suicidal behavior: What were the facts of the suicidal act itself? What were the proximate events leading to the crisis and the attempt? and What distal diathesis made the attempt possible? With some modifications, the ABCs can also assist in gathering information in patients who are evaluated for suicidal ideation or intent, not just for patients who are seen after a suicide attempt. The ABCs roughly follow the flow
of the clinical evaluation and can be regarded as a semistructured interview.

A—Acute assessment
B—Behavioral dissection
C—Crisis
D—Diagnostic 4 Ds (depressed, deranged, dysfunctional, delirious/diseased)
E—Ethanol and drugs
F—Family and personal suicide history
G—Guns
H—Homicide
I—Infanticide

Briefly, the A for acute assessment is a reminder that every patient needs to have an initial assessment of acute medical and acute safety issues following a suicide attempt: Were delayed toxic effects from an overdose considered? Were all necessary laboratory examinations conducted? Does the patient need cardiac monitoring? Were possible weapons removed? The B stands for behavioral dissection, using the interviewing technique of behavioral analysis developed by Pascal. This technique involves asking specific questions to ensure that all details of the attempt are inquired about (“What exactly did you do next?” “Did you hold the gun to your head?”), and it establishes chronology and fact rather than vagueness and opinion. After the behavioral dissection, the C crisis that precipitated the suicide attempt needs to be understood. What life event (usually a form of loss, real or imagined, e.g., loss of an idea; physical, social, or psychological loss) pushed the patient over the edge?

If possible, a psychiatric diagnosis is made. The 4 Ds signify specific and potentially treatable psychiatric conditions that must be ruled in or out: Is the patient depressed (including feeling anxious as an important affective state), deranged (psychotic), dysfunctional (personality disordered, particularly of the emotionally unstable variety), or delirious/diseased (medical illness including pain)? The importance of intoxication or withdrawal from (E) ethanol or drugs is self-evident, since ethanol and drugs can be depressogenic and, importantly, can lower impulse control. The (F) family and personal suicide history taps into biology as yet another source of lowered impulse control. (G) Guns address the question of availability of means, since over 50% of suicides occur by gun shot. And finally, (H) homicide and (I) infanticide are reminders to inquire about others at risk such as in an extended suicide, including children.

Absent from the ABCs are age and other demographic risk factors. I subsume these risk factors under crisis, e.g., What is the (life) crisis for this elderly widowed man who lives alone? for this young woman with an inexorably deteriorating neurologic disorder? for this college student who finds himself at home after his first psychiatric hospitalization for psychosis? for this gay man who transitions from HIV disease to AIDS?

Any physician might be called upon to evaluate a patient after a suicide attempt—no specialty is exempt. The ABCs of suicide provide all physicians with a concise yet comprehensive way of inquiring about what happened and why it happened. The ABCs help to document the assessment and to develop a treatment plan that is based on the specific circumstances of the patient. In that way, the ABCs help to provide competent clinical care and to address risk management concerns in potentially treacherous medico-legal situations.

Dr. Freudenreich has received grant/research support from Pfizer.

REFERENCES


Oliver Freudenreich, M.D.
MGH Schizophrenia Program
Massachusetts General Hospital
Boston, Massachusetts