ELSEVIER

RESPONSE

The focus of the papers by Katz et al. (1987, 1991) was identifying the time of onset of antidepressant effect. Lacking a placebo group, improvement was contrasted at week 1 and 2 for eventual responders versus non-responders at four weeks; differences were detected by week 1 and attributed to specific drug effects. Does this observation establish specific antidepressant drug effects occur at week 1? We believe this paradigm incapable of achieving its intended purpose since regardless of severity, patients receiving antidepressants also improve because of spontaneous remission and placebo effects, thus confounding this contrast. Katz et al. (1987, 1991) are aware of this since they state "From an experimental standpoint, it is difficult to separate which effects during drug treatment are due specifically to actions of the drug and which are due to non-specific factors such as placebo or spontaneous recovery. To resolve these factors sound experimental design would require a placebo group." (Katz et al. 1991, p. 600).

Katz et al. (1991) do not establish their procedure obviates the need for a placebo group. Their implicit assumption is that severely ill patients unchanged after a two week single blind placebo period would not exhibit placebo effects in the double blind phase. We could not find a single report of depressed patients void of placebo effect during a six week study. Within diagnoses, but across patient samples, a fluctuating level of placebo effects characterizes all psychiatric disorders independent of severity (Kane and Borenstein, 1996). The wide variability in placebo response rates necessitates a placebo control in all definitive psychopharmacologic trials.

Further, in the Katz et al. (1987) study even at baseline, the eventually recovered group of patients statistically differed from the eventually non-recovered group on severity of depressed mood, agitation, and cognitive impairment. Their use of the analysis of covariance to "equate" baseline differences is incorrect because enrollment in recovered and non-recovered groups is not random, and is subsequent to treatment (Table 1, Cohen, 1977, p. 301).

In order to show our objection was not purely academic, we demonstrated that the eventual six week responder versus non responder contrasts, early in treatment, were entirely consonant with placebo effects (Quitkin et al. 1996). Katz et al. (1997) in their letter suggest, differences in the effect sizes in the two sets of analysis vitiates our objections and attributes our smaller effect sizes to sample and dosage differences. This led us to question if our use of six week response may account for differences in effect size since Katz et al. (1987) used four week response. We have shown placebo responders on drug have an earlier onset than medication responders. (Quitkin et al. 1987). At week 4 (versus week 6), we hypothesized the responder group would contain a higher proportion of improvement attributable to placebo effects. In fact, using week 4 ultimate responders, our effect sizes (see Tables 1 and 2), are entirely comparable to those of Katz et al. (1987). Further, Croughan et al. (1988), examining the same data set as Katz et al. (1987) suggest there was little early observable clinical change, argued that "none of them responded to active treatment in the first week of drug therapy" suggesting no changes "were visible to clinical observers" (using a criterion of HAM D of 10). The importance of distinguishing early heuristically relevant change from clinically relevant change is discussed below.

Table 1. Recovered versus not Recovered at Week 4 (Using CGI), Proportion Rated Recovered at Week 1 Percent (%), Phi[a] (ϕ), and Effect Size (h)

	Recovered at Week 4 % and ϕ at Week 1	Not Recovered at Week 4 % and ϕ at Week 1	Effect Size (h) at Week 1 (Recovered versus not Recovered)
Placebo	26% (12/47) = 1.1	7% (10/154) = .54	1.1 – .54 = .56
Drug	15% (23/154) = .79	8% (18/238) = .57	.79 – .57 = .22

[a]The percentage and its arcsin conversion to ϕ (Phi) is presented. Arcsin transformation is used to correct for differences in detectability of proportions when calculating effect sizes (h); i.e., 10% to 0% is not equal to 50% – 40% (Cohen, 1977).

NEUROPSYCHOPHARMACOLOGY 1997–VOL. 17, NO. 2

Published by Elsevier Science Inc.
655 Avenue of the Americas, New York, NY 10010

0893-133X/97/$17.00
PII S0893-133X(97)00037-7

Table 2. Recovered versus not Recovered at Week 4 (Using CGI), Proportion Rated Recovered at Week 2 Percent (%), Phi[a] (ϕ), and Effect Size (h)

	Recovered at Week 4 % and ϕ at Week 2	Not Recovered at Week 4 % and ϕ at Week 2	Effect Size (h) at Week 2 (Recovered versus not Recovered)
Placebo	45% (21/47) = 1.5	7% (10/154) = .54	1.5 – .54 = .96
Drug	38% (58/154) = 1.3	13% (30/238) = .74	1.3 – .74 = .56

[a]The percentage and its arcsin conversion to ϕ (Phi) is presented. Arcsin transformation is used to correct for differences in detectability of proportions when calculating effect sizes (h); i.e., 10% to 0% is not equal to 50% – 40% (Cohen, 1977).

Katz et al. (1991) also report that drug plasma levels correlate with response at weeks 1, 2, and 3, supporting the assertion that the observed changes are attributable to true drug effect. If the total sample is examined, which is the usual plasma level study procedure, at week 1 amitriptyline concentration is correlated with "distressed expression", and nortriptyline to "increased agitation" and decreased "interpersonal sensitivity." At week 2, there is an "association between amitriptyline and a reduction in general psychopathology, and no other specific correlations". At week 3, there was a significant association between reduction of sleep disorder and plasma concentration of amitriptyline and nortriptyline, but no other significant relationships were found. These are meager findings that are not clearly related to antidepressant effect (as opposed to soporific effect).

Without any *a priori* hypothesis, Katz et al. (1991) also correlated amitriptyline and nortriptyline blood level with a variety of behavioral measures within responders. In Table 4 of Katz et al. (1991), data on plasma level for amitriptyline and nortriptyline are presented at week 2 and 3. In the text, correlations for responders at week 1 are discussed. Analyses using eleven measures on three occasions for amitriptyline and nortriptyline were conducted, yielding 66 possible correlations. While some measures are significant for amitriptyline and nortriptyline at weeks 1, 2, or 3, only somatization and sleep provide a consistent association on all three occasions, suggesting random fluctuations. Further, correlations for combined amitriptyline and nortriptyline (the usual standard), as well as imipramine or desipramine levels are not presented for responders or the total sample.

Kocsis et al. (1986) in another report on the same data examined amitriptyline, nortriptyline, imipramine, and desipramine levels and their combinations, and relate it with the outcome at week 4. No evidence of either linear or curvilinear relationships between plasma total concentration and improvement was found, suggesting that sparse, early, and largely inconsistent correlations have little relevance to specific antidepressant effect.

Katz et al. (1997) imply that Tollefson et al. (1994) supports their position. In discussing this issue, the context of the investigation requires definition, is it heuristic or clinical? If heuristic, defining the earliest sign of specific antidepressant effect occurs, is informative. To detect small, specific, behavioral effects, power requirements indicate that a large placebo-controlled study is appropriate. However, even highly significant differences in large samples resulting from a small effect size would not be clinically observable. Clinical relevance requires a categorical change observable within individual patients. If the interest is contrasting the clinical utility of treatments (drug versus drug, or drug versus psychotherapy), the total effect size and the proportion observed each week should be reported. Because even if one treatment had onset in week 1, whereas another had a later onset but a much larger effect by the 3rd and 4th week, the second treatment might be superior. Furthermore, there may be drug induced changes, such as sedation that are irrelevant to specific antidepressant effects. Therefore, the dependent variable must specifically measure change in depression.

Tollefson et al. (1994) defined patients with a 50% reduction in HAM–D as responders and remitters achieving a HAM–D of less than 8. There was no difference at the end of one week in the proportion of responders on drug and placebo, in spite of a sample of 1447 patients. With a sample number (N) this size, any specific effect must be small to be indiscernible. At the end of two weeks, the difference in the proportion of responders was 19% versus 11%, respectively for drug and placebo, an effect size of 0.19. There was no difference at week 3 in remitted patients on drug versus placebo. This hardly supports the notion that improvement attributable to drug can be clinically observed in the first 1–2 weeks. Effect sizes of .2 or less are unlikely to be observed clinically (Cohen 1977). We do not see how this allows the clinician early in treatment to "adjust his therapeutic approach accordingly." (Katz et al. 1997).

Frederic M. Quitkin, M.D.
Patrick J. McGrath, M.D.
Jonathan W. Stewart, M.D.
Bonnie P. Taylor, M.A.
Donald F. Klein, M.D.
New York State Psychiatric Institute
New York, New York

REFERENCES

Cohen J. (1977): Statistical Power Analysis for the Behavioral Sciences. New York and London, Academic Press

Croughan JL, Secunda SK, Katz MM, Robins E, Mendels J, Swann A, Harris–Larkin B (1988): Sociodemographic and prior clinical course characteristics associated with treatment response in depressed patients. J Psychiat Res 22(3):227–237

Kane J, Borenstein M (1996): Use of placebo control in psychiatric research, paper presented at CINP July, 1996 meeting. Melbourne, Australia

Katz MM, Koslow SH, Maas JW, Frazer A, Kocsis J, Secunda S, Bowden CL, Casper R, (1991): Identifying the specific clinical action of amitriptyline: Interrelationships of behavior, affect, and plasma levels in depression. Psychol Med 21:599–611

Katz MM, Koslow SH, Maas JW, Frazer A, Bowden CL, Casper R, Croughan J, Kocsis J, Redmond E Jr (1987): The timing, specificity and clinical prediction of tricyclic drug effects in depression. Psychol Med 17:297–309

Kocsis JH, Hanin I, Bowden C, Brunswick D (1986): Imipramine and amitriptyline plasma concentrations and clinical response in major depression. Br J Psychiatr 148:52–57

Quitkin FM, Rabkin JG, Markowitz JM, Stewart JW, McGrath PJ, Harrison W (1987): Use of pattern analysis to identify true drug response: A replication. Arch Gen Psychiatr 44:259–264

Quitkin FM, McGrath PJ, Stewart JW, Taylor BP, Klein DF (1996): Can the effects of antidepressants be observed in the first two weeks of treatment? Neuropsychopharmacology 15:390–394

Tollefson TG, Holman SL (1994): How long to onset of antidepressant action: a meta-analysis of patients treated with fluoxetine or placebo. Int Clin Psychopharmacol 9:245–250