

Thomas A. Ban: Neuropsychopharmacology in historical perspective. Education in
the field in the post-psychopharmacology era
Collated 16

Martin M. Katz: Onset of clinical effects of antidepressants
An interaction between Martin M. Katz and Donald F. Klein
with the participation of
Leslie C. Morey, Carlos Morra and Elemer Szabadi

Martin M. Katz – Donald F. Klein

Katz	January 9, 2014	essay
Klein	February 20, 2014	comment on Katz's essay
Katz	February 26, 2015	reply to Klein's comment
Klein	March 12, 2015	response to Katz's reply
Katz	April 2, 2015	response to Klein's response
Klein	July 16, 2015	response (2) to Katz's response
Katz	October 1, 2015	response (2) to Klein's response (2)

Lesley C. Morey – Martin M. Katz – Donald F. Klein

Morey	December 17, 2015	comment on Katz and Klein interaction
Katz	January 14, 2016	reply to Morey's comment
Klein	March 31, 2016	reply to Morey's comment
Katz	June 9, 2016	Katz's comment on Klein's reply to Morey's comment

Carlos Morra – Martin M. Katz – Donald F. Klein

Morra	May 28, 2015	comment on interaction between Klein and Katz
Katz	November 19, 2015	reply to Morra's comment
Morra	August 11, 2016	response to Katz's reply
Klein	January 21, 2016	comment on Katz's reply to Morra's comment
Katz	March 17, 2016	reply to Klein's comment on Katz's reply
Klein	March 24, 2016	corrected comment on Katz's reply to Morra's comment
Katz	July 14, 2016	reply to Klein's corrected comment on Katz's reply to Morra's comment
Klein	September 8, 2016	Klein's reply to Katz's comment on Klein's reply to Morra's comment

Carlos Morra – Donald F. Klein – Martin M. Katz

Klein	July 23, 2015	reply to Morra's comment
Morra	April 14, 2016	response to Klein's reply to his (Morra's) comment
Klein	July 7, 2016	response to Morra's response to Klein's reply
Morra	July 21, 2016	response 2 to Klein's response to Morra's response
Katz	October 15, 2015	comment on Klein's reply to Morra's comment
Klein	September 8, 2016	reply to Katz's comment on Klein's reply to Morra's comment
Katz	May 26, 2016	comment on Morra's response to Klein's reply to Morra's comment

Elemer Szabadi – Martin M. Katz – Donald F. Klein.

Szabadi	August 13, 2015	comment on interaction between Klein and Katz
Katz	October 29, 2015	reply to Szabadi
Klein	October 22, 2015	reply to Szabadi
Katz	October 29, 2015	reply to Szabadi

Martin M. Katz - Donald F. Klein

Martin M. Katz: Onset of Clinical Action of Antidepressants

Time of onset of clinical actions induced by antidepressants (ADs) is critical for uncovering basic mechanisms underlying their efficacy, for developing further understanding of the nature of the depressive disorder and for predicting early in treatment, whether a drug is likely to be effective. The criticality of the onset issue has been recognized since the discovery of the new drugs. It was initially observed by Kuhn (1958) that the clinical effect in most responsive patients occurred within the first week. The controversy was then, ignited by the Quitkin et al. (1984) study showing that clinical actions of the antidepressants “lag” several weeks behind the drug’s initial effects on central neurotransmitter systems. The latter study resulted in the “onset lag” becoming a commonly accepted “textbook” notion. Conversely, the National Institute of Mental Health (NIMH) Collaborative Depression Study group (CDS) reported in Katz et al. (1987), based on a large sample of severely depressed, hospitalized patients, that in treatment–responders, significant changes in major components of the disorder occurred within the first two weeks. Neither of the two studies was aimed directly at the “onset” issue; the results on onset were the product of secondary analyses. Neither study was, therefore, able to provide a definitive answer to the question of clinical-onset. What the studies did accomplish, however, was to highlight the onset question, critical to determining sequence of drug actions and to uncovering relationships between drug-induced neurochemical and clinical actions. At the practical level, knowledge of timing also determines when the clinician can expect to see the first drug-induced changes, and whether the presence or absence of early changes can predict the nature of the patients’ clinical response to drug treatment.

Following these early reports, technical papers aimed at the methodology required to achieve definitive answers on timing appeared, and a series of independent meta-analyses targeting the problem in large drug trials, were conducted. A body of literature on the issue has been developed since 1990 that many now believe have resolved the issue.

An abbreviated set of references, including papers which analyze this area of the literature, and which report the more definitive results from the meta-analytic studies, is listed below. The

list includes the published earlier exchange on the two conflicting views in the journal *Neuropsychopharmacology*.

The general consensus as it exists today can be summarized in the following statements drawn from several of these recent publications:

1. “One-third of the total (clinical) effect of selective serotonin re-uptake inhibitors (SSRIs) after six weeks of treatment is seen in the first week” (Taylor et al. 2005, based on literature review and meta-analyses).
2. “Among responders the onset of improvement occurs in more than 70% of cases within the first three weeks of treatment with an AD” (Stassen et al. 1997, based on analysis from a multihospital study and survey of results).
3. “Drug specific types of behavioral response in the first one or two weeks of treatment with desipramine or paroxetine are highly predictive of six-week outcome” (Katz et al. 2004, based on drug-placebo comparison study).
4. Absence of behavioral changes during first two to three weeks indicates little chance of positive response at outcome (Szegedi et al. 2009; Katz et al. 2011; Stassen et al. 1997, based on finding that >90% of patients who show no improvement during the first two weeks, show non-response at outcome)

Relevant References:

I. Technical: How to measure onset

Katz MM, Koslow SH, Fraser A. Onset of antidepressant activity: reexamining the structure of depression and multiple actions of drugs. *Depression and anxiety*, 1997;4:257-67.

Laska EM, Siegel C. Characterizing onset in psychopharmacological clinical trials. *Psychopharmacolog Bull*, 1995;31:29-35.

Leon AC, Blier P, Culpepper L, Gorman JM, Hirschfeld RM, Nierenberg AA, Roose SP, Rosenbaum JF, Stahl SM, Trivedi MH. An ideal trial to test differential onset of antidepressant effect. *J Clin Psychiatry*, 2001; 62:34-6.

Stassen HH, Angst J, Delini-Stula A. Time course of improvement under antidepressant treatment: a survival-analytic approach. *Eur Neuropsychopharmacol*. 1993;3:127-35.

II. Clinical Studies and Reviews:

Kuhn R. The treatment of depressive states with G22355 (imipramine hydrochloride). *American Journal of Psychiatry*, 1958;115:459-64.

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

Katz MM, Tekell J, Bowden CL, Brannan S, Houston JP, Berman N, Frazer A. Onset and early behavioral effects of pharmacologically different antidepressants and placebo in depression. *Neuropsychopharmacology*, 2004;29:566-79.

Katz MM, Berman N, Bowden CL, Frazer A. The componential approach enhances the effectiveness of 2-week trials of new antidepressants. *J Clin Psychopharmacology*, 2011;37:193-218.

Machado-Vieira R, Luckinbaugh DA, Manji HK, Zarate CA. Rapid onset of antidepressant action: a new paradigm in the research and treatment of major depressive disorder. *J Clin Psychiatry*, 2008;69:946-58.

Posternak MA, Zimmerman MD. Is there a delay in the antidepressant effect? A meta-analysis. *J Clin Psychiatry*, 2005;66:148-58.

Papakostas GI, Perlis RH, Scalia MJ. A meta-analysis of early sustained response rates between antidepressants and placebo for the treatment of major depressive disorder. *J Clinical Psychopharmacol*. 2006;26:56-60.

Quitkin F M, Rabkin JG, Ross D, Stewart JW. Identification of true drug response to antidepressant. *Arch Gen Psychiatry*, 1984;41:782-6.

Robinson DS. Antidepressant drugs: Evidence for onset of therapeutic effect. *Primary Psychiatry*, 2007;14:21-3.

Taylor MJ, Freemantle N, Geddes JR, Bhagwagar Z. Early onset of SSRI antidepressants: Systematic review and meta-analysis, 2006;63:1217-23.

Stassen HH, Angst J, Delini-Stula A. Delayed onset of action of antidepressant drugs? Survey of recent results. *Eur Psychiatry*, 1997;12:166-76.

Szegedi A, Jansen WT, van Wugenburg AP. Early improvement in the first two weeks as predictors of treatment outcome in patients with major depressive disorder: a meta-analysis including 6,562 patients. *J Clin Psychiatry*, 2009;70:344-53.

III. Earlier Controversy: Conflicting Views on the Evidence

Can the effects of antidepressants be observed in the first two weeks of treatment?

Quitkin et al. Neuropsychopharmacology, 1996;15:390-4.

Katz MM et al. Neuropsychopharmacology, 1997;17:110-1.

January 9, 2014

Donald F. Klein's Comment

Martin Katz in the Controversy section of INHN stated that Fred Quitkin started a controversy by claiming that the clinical actions of antidepressants lag several weeks behind the drug's initial effects on the CNS. It might be helpful to recall the background of that particular 1984 study (Quitkin FM, Rabkin JG, Ross D, Stewart JW. Identification of true drug response to antidepressants. Use of pattern analysis. Arch Gen Psych 1984; 44: 259-64).

Each of 185 patients were administered weekly Global Improvement Scores while receiving placebo or medication double-blind. The question was whether there was some trajectory peculiar to drug treatment. It seemed simplest to dichotomize these scores, to either zeroes or ones. This meant frank remission or unimportant symptomatology without dysfunction was rated zero, and all other scores, rated one. Since there were five weeks of scoring during treatment, there were 32 conceivable patterns of consecutive ones and zeros.

Remarkably, certain patterns never appeared in the placebo group, but were relatively common during drug treatment. Further, they were markedly similar to each other. Within this group, each string was initiated by a zero, at any week, and was also rated zero for all subsequent weeks. The only exception was that this series never started at week one. There were patients whose initial score was zero, but these inevitably went downhill. We were quite pleased with this result since it confirmed our clinical impressions - especially the persistence of benefit.

Note that this is overstated by Katz, who cites, the "Quitkin et al. (1984) study showing that clinical actions of the antidepressants 'lag' several weeks behind the drug's initial effects on central neurotransmitter systems."

What is the subtext here? We guess that a certain model of pathophysiology and repair of neural functioning during depression is at stake. It was discovered that imipramine rapidly blocks the synaptic reuptake mechanism, delaying the exit of the neurotransmitters from the synaptic cleft.

It was initially assumed that this excessive synaptic stimulus would relieve the brain of the functional decrement that underlay depression. (It was never quite clear why the excess neurotransmitter did not lead to receptor desensitization. Inhibitory afferent autoceptors were not part of the machinery yet). It followed that the antidepressant effect should be very rapid since the hypothesis was that the neurotransmitter deficit was directly manifested as depression. This fit well with simplistic advertising that strongly implied that a depression was due to a neurotransmitter deficit so that you had norepinephrine and serotonin depressions, as well as norepinephrine and serotonergic therapies.

If there was a lag between drug administration and antidepressant effect, it conceptually demoted these neurotransmitters into being, at best, the first domino initiating a complex cascade involving who knows what.

Strangely this heuristic question relating the initial impact of medication to eventual clinical repair has gotten twisted into studies arguing for a quick medication effect, as if that corrected some misapprehension Quitkin had generated. After all, Quitkin had not claimed a “several week delay” was necessary.

The supposed opinion difference regarding onset gap has been further twisted to rest upon whether it can be shown that there is a statistically significant drug superiority to placebo within the first week.

Now we can get into a highly technical, mathematical exposition concerning the power that would allow detection of a small drug - placebo difference at week one, the necessity for a multisite study, the increase in diagnostic error, deterioration of reliability, etc. However, fortunately, all of that is completely unnecessary .

Even if it were true, that under some circumstances medication was substantially better than placebo, during the first week of administration, it is certainly not the usual situation.

The majority of drug responsive patients, even in all the cited studies that Katz believes affirms his position, achieve remission after several weeks - just as Quitkin et al. affirmed. The heuristic question has been answered - the gap exists. The immediate effects of antidepressants on neurons are insufficient for understanding the process of recovery.

Perhaps there are practical issues that hinge on whether there is an early therapeutic response or not. Katz has made three suggestions that depend upon the supposed ability of a small improvement during the first two weeks being highly related to a good eventual outcome. Further,

if during the first two weeks there is not even slight evidence of improvement, it is extremely unlikely that this treatment will work. Therefore, switching treatments is a real option early during an unsatisfactory treatment. Also, since early response is so closely tied to the eventual outcome, there is no reason why the clinical trial cannot be radically shortened, to say two- and one-half weeks, and enormous savings incurred.

One strange aspect of these studies is their lack of inclusion of a placebo group in these analyses, which are essentially within drug group predictive analyses. This, of course is highly problematic. Further, the numerical basis for many claims in this area is often obscure. We are fortunate that Katz has provided relevant data in: Katz MM, Berman N, Bowden CL, Frazer A. The componential approach enhances the effectiveness of 2-week trials of new antidepressants (J Clin Psychopharmacology 2011; 37: 193-218).

TABLE 1. Predictive Accuracy of 2-Week Improvement Values: Percentage of Correct Component Predictions at Treatment Outcome (Modified).

	Desipramine (n = 26)			Paroxetine (n =24)		
	TP	TN	C	TP	TN	C
Depressed mood	62	100	68	58	89	71
Anxiety	55	33	52	25	100	43
Motor retardation	76	100	80	40	64	52
Hostility	53	63	56	46	88	62
Anx-Agit	69	67	68	33	91	67
DM-MR	67	100	72	64	80	52
Hamilton total	64	50	62	64	90	75

Table 1 is a direct transcript of the text, removing significance signs. TP is the proportion of $\geq .5$ outcomes correctly called. TN is the proportion of $< .5$ outcomes correctly called. $C = TP + TN$.

From this data the basic 2X2 table relating early small gains, $\geq 20\%$, to later findings of substantial, $\geq 50\%$, gain can be definitively reconstructed. See Table 2 (without rounding to nearest integer).

PAROXETINE	PREDICTION UNDERSHOOT PREDICTION OVER SHOOT			DMI	PREDICTION UNDERSHOOT PREDICTION OVER SHOOT								
	>=50%	<50%	POS OUTCOMES		>=50%	<50%	POS OUTCOMES	NEG OUTCOMES					
DEP MOOD	>=20%	8.08	1.11	9.19	0.34	0.47	DEP MOOD	>=20%	13.57	0.00	13.57	0.38	2.03
	<20%	5.85	8.96	14.81				<20%	8.32	4.11	12.43		
		13.94	10.06	24.00					21.89	4.11	26.00		
ANX	>=20%	4.56	0.00	4.56	0.75	2.38	ANX	>=20%	12.35	2.38	14.73	0.34	2.18
	<20%	13.68	5.76	19.44				<20%	10.10	1.17	11.27		
		18.24	5.76	24.00					22.45	3.55	26.00		
MOT RET	>=20%	4.80	4.32	9.12	0.24	0.24	MOT RET	>=20%	16.47	0.00	16.47	0.24	1.20
	<20%	7.20	7.68	14.88				<20%	5.20	4.33	9.53		
		12.00	12.00	24.00					21.67	4.33	26.00		
HOST	>=20%	6.52	1.18	7.70	0.46	0.66	HOST	>=20%	9.65	2.89	12.53	0.31	0.73
	<20%	7.65	8.65	16.30				<20%	8.55	4.91	13.47		
		14.17	9.83	24.00					18.20	7.80	26.00		
ANX AGIT	>=20%	3.28	1.27	4.54	0.54	0.38	ANX AGIT	>=20%	8.97	4.29	13.26	-0.02	-0.02
	<20%	6.65	12.80	19.46				<20%	4.03	8.71	12.74		
		9.93	14.07	24.00					13.00	13.00	26.00		
DM-R	>=20%	8.64	2.10	10.74	0.20	0.26	DM-R	>=20%	14.78	0.00	14.78	0.33	1.85
	<20%	4.86	8.40	13.26				<20%	7.28	3.94	11.22		
		13.50	10.50	24.00					22.06	3.94	26.00		
HAM TOT	>=20%	8.86	1.02	9.88	0.29	0.39	HAM TOT	>=20%	14.26	1.86	16.12	0.28	1.66
	<20%	4.98	9.14	14.12				<20%	8.02	1.86	9.88		
		13.85	10.15	24.00					22.29	3.71	26.00		

Several findings then appear. The positive outcome proportions predicted are substantially less than those obtained. The proportion of subjects who do well is usually about 30% greater than the proportion predicted to do well. The claim that those who do well initially, will also do well later, is true but misleading. The undershoot invalidates the claim for a short clinical trial since the drug would be undervalued.

The claim that if the patient does poorly on all initial variables, they will do poorly later, cannot be evaluated since only individual variables are available. However, for these individual variables, the chance of doing well despite poor initial performance is substantial for paroxetine but looks even better for DMI. Here the drugs are predictively devalued which casts doubt on the value of early treatment change, given initial disappointment. It would also seem likely that those who do poorly on all initial variables are only a small proportion of the sample and may also be quite atypical on other grounds.

It should be noted that a focus on the immediate effects of medication on neurotransmitters yields a supposedly promising, fairly narrow, pathway to the development of agents that will improve the process and thereby act therapeutically. On the other hand, if one has to deal with a complex cascade, our theory of depression becomes quite obscure and the directions that one can take in pursuing remediation appear all too many. As the development of antidepressants has been almost exclusively a matter of serendipity, it is plain that understanding pathophysiology and repair is still well beyond us. It seems unlikely that translational thrashing about, using limited current knowledge, will prove profitable.

To sum up, Katz has made a heuristic and several practical suggestions relating to clinical trials. These suggestions are not supported by his data. Similar re-analyses of data, whose current analyses claim to support Katz's views, would be very worthwhile.

February 20, 2014

Martin M. Katz's reply to Donald F. Klein's comment

Don Klein begins his comment by recalling the background of the 1984 Quitkin et al. study that concluded that antidepressants (ADs) lagged several weeks behind the initial "clinical" effects on the CNS.

I am not persuaded that his description of the fine points of that study changes any of the facts cited in my essay. The Quitkin study was innovative in its approach to distinguishing specific drug responders from placebo responders. I also, with other interpreters of their results, noted that the Quitkin study was not designed to determine the time of onset AD clinical actions. The results could not be generalized to inform about onset because of the modest dosage of imipramine used to treat study patients, the graduated dosage schedule applied so that even this modest dosage level was not achieved until the end of the first two weeks and the insensitivity of the measures used to detect clinical effects, if they existed, during this period. I am not critical of the study otherwise, only that it has little to tell us about onset of AD clinical actions. On the issue of when "full response" to antidepressants is achieved, all would agree with Klein that that does not usually occur before several weeks of treatment. Nevertheless, the question here is not about full response, but whether the drug is inducing significant clinical actions early, within the first two weeks, of the course of treatment. Here the jury is not out; several independent multisite, large sample studies have affirmed that such actions do occur, sometimes as early as the first week. All also agree about the importance of such findings toward understanding the neurobehavioral mechanisms underlying the efficacy of antidepressant drugs and the practical implications of such findings for the future design of clinical trials and for treatment practice.

The other studies referred to by Klein, which also included placebo controls (Stassen et al. 1996, Szegedi et al. 2009, Katz et al. 2004), sought to determine onset by various techniques, one of which involved testing significance of difference in changes at one and two weeks of treatment

between drug and placebo, another by comparing the number of patients who showed a $\geq 20\%$ decrease in the Ham-D total score (a decrease validated by Stassen et al. to be clinically significant) at the early time points between drug and placebo.

It was further demonstrated in these studies that this amount of “early improvement,” i.e., $\geq 20\%$ decrease, was clinically significant in that it could predict that 70% of patients showing this early improvement would go on to respond at 6 or 8 weeks to the experimental treatment; even more telling, that less than 10% of patients who did not show this early improvement, did not change course and respond at outcome.

Klein’s concern is that these findings could not be applied to individual patients (he apparently believes despite the finding of drug-placebo differences, that is “not the usual situation”). The facts evidenced from the studies referenced above indicate that on the contrary, on average, patients will respond in the manner described in these carefully designed studies.

We do make a case as he notes, that these results support shortening the standard clinical trial, and that our conclusions are strongly supported by the accumulated evidence on prediction in this area. Klein turns to our 2011 paper and analysis in the second part of his comments, to dispute the interpretation of the results presented. I am, however, unclear about how he has reanalysed the data from our table in that paper, what procedures he actually carried out, so am unable to respond intelligibly to this section of his comments.

February 26, 2015

Donald F. Klein’s response to Martin M. Katz’s reply

TABLE 1. Predictive Accuracy of 2-Week Improvement Values: Percentage of Correct Component Predictions at Treatment Outcome (Modified).

	Desipramine (n = 26)			Paroxetine (n =24)		
	TP	TN	C	TP	TN	C
Depressed mood	62	100	68	58	89	71
Anxiety	55	33	52	25	100	43
Motor retardation	76	100	80	40	64	52

Hostility	53	63	56	46	88	62
Anx-Agit	69	67	68	33	91	67
DM-MR	67	100	72	64	80	52
Hamilton total	64	50	62	64	90	75

Table 1 is a direct transcript of the text, removing significance signs. TP is the proportion of $\geq .5$ outcomes correctly called. TN is the proportion of $< .5$ outcomes correctly called. $C = TP + TN$.

Table 1 reproduces Katz's findings as presented in my Comment on "Onset of antidepressant action."

It was a mistake not to spell out the simple algebra that allows the reconstruction of the original 2x2 data tables, from the indices in Table 1.

The predictor is the group of those subjects who had at least 20% improvement on a specific variable vs the others, during the first two weeks of the trial. The outcome variable is the group of those subjects who had at least 50% improvement on that variable at trial end vs the others.

The cells of the 2X2 are usually designated

A	B
C	D

Adding up to N. The number called positive at outcome is (A+C). The number called positive at 2 weeks is A+B. Since we don't know the Total Number positive at outcome (=True positives +False Positives) we stipulate it as N-X therefore the number called negative at outcome is X.

That allows the simple formula: $TP*(N-X) + TN*X = C*N$. Note the variable C is not the cell label "C" but rather, TP + TN. Entering the given values determines X, which allows filling in the entire 2X2.

$$A = TP*(N-X)$$

$$D = TN*X$$

$$C = (N-X) - (TP*(N-X))$$

$$B = X - TN*X$$

	$\geq .5$	$< .5$
$\geq .2$	#TP=TP* (26-x)	
$< .2$		#TN=TN*x
	26-x	X

The results are not exact integers due to rounding in Katz's values. These figures are displayed in Table Two of my initial comment.

The important issue is that they support every negation of all of Katz's hypotheses, as previously detailed. Since my analysis is now understandable. Marty can address the fact that his findings do not support his theories.

March 12, 2015

Martin M. Katz's response to Donald F. Klein's response

This is a follow up to my earlier response to Don Klein's reanalysis of our data from the 2011 paper on predicting positive outcome based on early improvement.

I wish to note at the beginning that the 2011 study (Katz et al.) represented an attempt to demonstrate that prediction at two weeks is possible and that adopting the approach we outlined in that paper would be useful in any effort to shorten the length of clinical trials. I refer to the study as an example and acknowledge that it is based on a relatively small patient sample; so, it is obviously not the sole foundation for major prediction in this area. This small study, as noted in earlier papers, should be followed by a prospective study with a large and representative patient sample to establish the validity and practicality of this approach and to establish the notion that a two-week trial would be sufficient to determine whether a new, putative antidepressant will be efficacious.

My essay on onset of antidepressant action mainly refers to findings from independent and large sample multisite studies (e.g., Stassen et al. 1996, Szegedi et al. 2009, Katz et al. 2004) that are relatively definitive in establishing that actions of efficacious drugs begin as early as 1 week. This is in contrast to the earlier textbook notion that the onset of clinical effects of these drugs lag several weeks after the almost immediate drug-induced actions on central neural transmitter systems. As an example of the process of prediction, the 2011 paper provides a test trial to demonstrate that significant prediction can be established at the two-week point. To accomplish that goal, we illustrated in Table 1 of that paper that combining the true positive predictors (TP) for DMI with the true negative predictors (TN) divided by N equals the percent of true predictors (C). Klein apparently used the wrong definition of C in his equations and although the algebra is correct, the results lack meaning.

Also, under DMI in the first line in Table 1, by adding TP and TN for “depressed mood”, he gets a total of 1.62, which means more patients than we started with. If you try using 0.68 for C, answers come out closer to Klein’s, but are not correct, e.g., if TN is equal to 1 then the cell for $EI < .20$ and outcome $> .50$ would be zero. So, I am still not clear what he means by overshoot or undershoot.

Klein may have a point about the two-week results sometimes not being an accurate predictor. We have made the point in earlier papers that all could be clarified if the larger prospective study is eventually conducted. In sum, I think that Klein has not fully understood the results projected in Table 1, leading to some miscalculations in his algebra. That aside, however, his analysis does not negate our conclusions. More important, as noted, the 2011 study represents only a small aspect of the entire base of information that allows us to conclude that onset of

efficacious agents occurs during the 1st two weeks. The test study merely represents an attempt to demonstrate how utilizing early improvement as a predictor and shortening the trial can lead to clinical benefits for the patient in reduced exposure to ineffective agents, and major cost reduction for the drug companies that develop and evaluate the new drugs. I appreciate Dr. Klein's study of our work and that of the others who have carried out the multisite clinical trials on which these conclusions are based, but respectfully submit that his analysis of our small study does not in any way negate these important conclusions.

References:

Katz MM, Berman N, Bowden CL, Frazer A. The componential approach enhances the effectiveness of 2-week trials of new antidepressants. *J Clin Psychopharmacology*, 2011;37:193-218.

Katz MM, Tekell J, Bowden CL, Brannan S, Houston JP, Berman N, Frazer A. Onset and early behavioral effects of pharmacologically different antidepressants and placebo in depression. *Neuropsychopharmacology*, 2004;29:566-79.

Stassen HH, Angst J, Delini-Stula A. Delayed onset of action of antidepressant drugs? Survey of recent results. *Eur Psychiatry*, 1997;12:166-76.

Szegedi A, Jansen WT, van Wugenburg AP. Early improvement in the first 2 weeks as a predictor of treatment outcome in patients with major depressive disorder: a meta-analysis including 6562 patients. *J Clin Psychiatry*, 2009;70(3):344-53.

April 2, 2015

Donald F. Klein's response (2) to Martin M. Katz's response

Katz points out limitations of his 2011 study: "...an attempt to demonstrate that prediction at two weeks is possible and that adopting the approach we outlined in that paper would be useful in any effort to shorten the length of clinical trials. This small study, as noted in earlier papers, should be followed by a prospective study with a large and representative patient sample to establish the validity... that a two-week trial would be sufficient to determine whether a new, putative antidepressant will be efficacious... mainly refers to findings from independent and large

sample multisite studies (e.g., Stassen et al. 1996; Szegedi et al. 2009; Katz et al. 2004) that are relatively definitive in establishing that actions of efficacious drugs begin as early as 1 week.”

This does not follow. What is the purpose of the small study that requires a large follow up if it has already been established by large studies?

Are these large studies? Unfortunately, only the insufficient abstracts of Szegedi could be retrieved. Katz (2004) had an N of 82, 12 dropped out after randomization. "... it was decided *a priori* that patients who did not complete at least 3 weeks of treatment would not generate useful data.” The remaining 70 were distributed into 3 groups, paroxetine, desmethylimipramine and placebo. Of the 29 in the DMI group, 3 dropped out by 2 weeks, of the 28 in the paroxetine group 4 dropped out by 3 weeks. In any case this is not a large study groups, paroxetine, desmethylimipramine and placebo

There is no simple listing of behavioral measures. I count 26 but this may be a substantial undercount. Similarly, the number, timing, and evaluator of the behavioral measures is not simply tabulated ,although very frequent. This affords an ample opportunity. The number of biochemical assessments is not stipulated. The analyses use Last Observation Carried Forward, a questionable practice. There is no mention of correction for multiple analyses. Therefore, the findings are not clearly distinguished from mere sampling variation.

The point of the study is detection of therapeutic onset. It was defined here as, “The 'median time of onset' was defined as the earliest time point at which 50% of patients changed a minimum of 20% on a given behavioral construct, a change that was then sustained throughout the course of treatment.”

This arbitrary measure does not derive from some statistical model of onset. Note that it yields a single index for each group since it depends on 50% of the group reaching the arbitrary 20% decrement that is sustained. The time of onset is likely to be variable among subjects. A definition of onset that can be individually applied would give some idea of the spread of onset times. Stassen (below) arbitrarily develops such a measure. I could not follow the analysis described for "Analysis of onset of 'therapeutic' action within each treatment group." It did not seem to yield an onset time, at least to me .Clarification may be helpful.

The analyses in Katz' latest INHN submission, used within drug comparisons of binary status measures. These are not demonstrations of drug effect, since they lack a placebo comparison.

In the section "Prediction of outcome, Logistic regression (Hosmer and Lemeshow 1989) was used to develop an algorithm for estimating the probability that a patient would recover by 6 weeks of treatment based on values on the behavioral constructs after 1 or 2 weeks of treatment. Different models of individual prediction were tested for each drug independently."

"We did not test models including variables that did not discriminate between recovered and non-recovered subjects at any of the early time points...The model and threshold that provided the best combination of sensitivity and specificity was then selected as the prediction model for recovery." This is exploratory work. That is justified but it should not be presented as definitive.

All of these shifting procedures plus the lack of correction for multiplicity of analyses of the same data set, leads to an analogy with the Texas Sharpshooter who carefully draws a target around each scattered bullet hole.

No doubt Katz was attempting to solve difficult problems by exploratory work.

Stassen (1996) states: "The sample consisted of moderately depressed male (n = 154) and female (n = 275) patients (aged 17-73), diagnosed according to DSM-III criteria for major depression. Of these, 120 were treated with oxaprotiline, 120 with amitriptyline and 189 with placebo. Efficacy criteria were Hamilton Depression (HAMD) and Anxiety (HAMA) and Zung Self-Rating scales. Up to eight ratings over a period of 40 days were available for analysis... the appropriate determination (is) of the time points at which the medication begins to clearly show a therapeutic effect in each individual patient... A solution to this problem is to define onset of improvement in each individual case on the basis of significantly reduced psychopathology scores relative to the corresponding baseline, that is, a reduction of - d% of baseline... Lacking appropriate a priori knowledge we unified as a tentative step all 429 cases (minus 17 cases due to insufficient data) to one single sample in order to get an estimate of the 'natural' variability of HAMD and HAMA scores over time ...It turned out that a relative change of 15-25% with respect to the corresponding initial values represents a suitable threshold for a reasonable definition of onset of improvement... we decided in favor of the 20% threshold."

Clearly the problem of therapeutic onset within individuals has not been solved. I could not find the demonstration of drug effects at 1 or 2 weeks that Katz refers to. The relevance of this paper to Marty's hypotheses is unclear.

Now to address the algebra! But why should we? I attempted to reconstruct the basic 2X2 tables relating early response to later response from the proffered indices referring to true

positives ,true negatives, overall correctness and sample size. Perhaps it was more complicated than I realized since the definitions given of the indices was complex and easily misunderstood. But Marty has that data. Surely the most easily understood, trenchant refutation would be the direct comparison of the actual 2X2 predictions with my reconstructions.

I also regret this data was not presented for they test the validity of Marty's hypotheses. He believes that important practical implications, such as shortening the length of clinical trials, follows. Certainly, presenting these 2X2's has more important implications than just refuting my analysis. However, fortunately, he still has that presentation opportunity

July 16, 2015

*Martin M. Katz's response (2) to Donald F. Klein's
Response (2)*

Don Klein's response raises several questions about my statements in the original essay that probably could have been answered with more careful reading of the referenced sources.

For example, regarding his first query about why the need for a large prospective study to confirm findings from our study reported in Katz et al. 2004, when such "large sample" studies of Szegedi et al. 2009 and Stassen et al. 1997 have already been conducted.

The answer is that the referenced studies used only the Ham-D as an outcome measure, whereas our 2004 study reports results with an array of dimensional measures of depression that reinforce the predictability of the early changes. Since definitive demonstration of one or more of these dimensions as a predictor increases the power and breadth of prediction, the results of the large sample study can be useful in determining the range of effects of which a new drug might be capable, if included in such studies.

Klein further questions whether Szegedi et al. (2009) and Stassen et al. (1997) are "large" studies, since he was only able to retrieve and study the abstracts. But even the Szegedi abstract (and title) retrieved from PubMed, notes that the sample size included 6,000+ patients and the Stassen abstract refers to some 1,500 patients, figures hard to ignore in reading the abstracts. Then, Klein diverts the discussion to our 2004 study, noting that the sample was 82 patients. He takes issue with the "multiple measures" and the lack of correction for multiple analyses. He has a point

here, but we were clear in the text to indicate that we confined the analyses to “specific hypotheses, focused on five constructs and one of the severity dimensions”, thus, limiting the number of analyses. Further, many of the primary findings had p-values less than the 0.05 level and were consistent across several methods of statistical analysis. So that if one had questions here, an even more conservative approach, such as accepting only those findings at the $p < 0.01$ level, findings that could clearly not be due to chance variations, would strongly reinforce the validity of the major study findings.

Regarding the algorithm for prediction referred to in the 2004 study, he questions the “definitiveness” of this finding. I do not see where we, at any point in the article, justified the algorithm as “definitive.” Our major point was that that was the best model that could be achieved by combining variables from a sample of this size. In view of the way it was applied, one would consider the algorithms, designed to provide the best combination of sensitivity and specificity, exploratory, in nature, another reason, to recommend a prospective large sample study on the prediction issue.

To achieve a reasonable estimate of when onset of clinical effects occurred in each drug and placebo group, we chose to apply the “median” time of onset approach, i.e., the time at which >50% of patients within a group, showed >20% change or significant improvement on that dimension. It seemed to us and to Stassen and Szegedi to be a highly defensible criterion for estimating the time at which a drug initiates significant improvement. Don Klein is welcome to differ with us on that, but it is not clear on what basis. The fact that individual patients will vary on this measure of onset is of course, obvious, but here we are simply seeking an adequate estimate of the time at which most patients (>50%) in this treatment group, show a significant amount of change; the median measure provided a simple and accurate estimate.

The “analysis of onset” paragraph on pg. 569 of Katz et al. (2004) referred to was to determine when improvement that leads to clinical response at outcome for the treatment-responder group, begins. The statistics were aimed at determining the initial (at two weeks) indicators for those behavioral variables that distinguished the responders from the non-responders to the specified treatment. Effect size was then used to see whether the behavioral changes at two weeks that distinguished the treatment responders, were not only statistically significant, but likely to be “visible” to the observer.

Klein ends his queries with a quote from Stassen et al. that clearly describes how their group arrived at > 20% change as a reasonable, statistically based definition of onset of improvement. After Klein claims not to have found the demonstration that I refer to as drug effects at 1 and 2 weeks (despite much of the original essay prepared on “onset” dedicated to establishing that the evidence was cumulative and quite compelling on this critical issue), he concludes that “clearly the problem of therapeutic onset has not been solved.” Maybe not completely, but we wonder: what is the basis, the studies that support his unacceptance of the voluminous evidence compiled over the last three decades that lead to the very logical, and well supported conclusion that the established antidepressant drugs begin their clinical effects within the first two weeks of treatment?

Regarding the table from our 2004 study, that Klein modified: that query was answered in my previous response.

References:

Katz MM, Tekell J, Bowden CL, Brannan S, Houston JP, Berman N, Frazer A. Onset and early behavioral effects of pharmacologically different antidepressants and placebo in depression. *Neuropsychopharmacology*, 2004;29:566-79.

Stassen HH, Angst J, Delini-Stula A. Delayed onset of action of antidepressant drugs? Survey of recent results. *Eur Psychiatry*, 1997;12:166-76.

Szegedi A, Jansen WT, van Wugenburg AP. Early improvement in the first two weeks as predictors of treatment outcome in patients with major depressive disorder: a meta-analysis including 6,562 patients. *J Clin Psychiatry*, 2009; 70:344-53.

October 1, 2015

Leslie C. Morey – Martin M. Katz – Donald F. Klein

Leslie C. Morey's comment

In following the interesting discussion between Dr. Klein and Dr. Katz with regard to the onset of antidepressant effects, there appears to be an opportunity to illustrate the implications of research results on clinical decision-making. In attempting to do so, I will use the 2x2 tables

provided by Dr. Katz from the Katz et al. (2011) paper to explore some of the concerns raised by Dr. Klein, and I believe that both Drs. Klein and Katz raise important points. Please note that this comment seeks to illustrate a point using the Katz data; I make no claims about interpreting the general trends in this research literature.

I believe that one of the central critiques from Dr. Klein is that the data from Katz et al. (2011) do not necessarily support a recommendation for early discontinuation of an antidepressant based upon a lack of an early response. Dr. Katz notes that there are clear data suggesting that active drug demonstrates superiority to placebo, at a group level, within the first week or two of treatment. However, Dr. Klein points out, accurately I believe, that these group differences do not necessarily support the premise of recommending early discontinuation for a particular patient. This is basically an issue of decision theory, into which we must use Bayesian concepts to assist our decision making. The most relevant concept for this particular issue is that of “negative predictive power” or NPP values—in other words, the probability that, if we have decided there will not be a treatment response based upon early indications, what is the probability that such a decision will be correct? I concur with Dr. Klein that the numbers presented in Katz et al. (2011), which highlight “Percentage of Correct Predictions at Treatment Outcome,” are examples of “sensitivity” values — which do not directly address the advisability of recommendations based upon early response. As noted by Dr. Klein, what is needed to determine these NPP values are the full 2x2 (early response by final response) results from that study, which Dr. Katz has provided in his most recent comment.

However, there are inconsistencies in Dr. Katz’s data in his 2x2 tables. In his first two tables (concerning Depressed Mood and Anxiety), he reports that 27 patients in his study demonstrated a 50% decrease on the HAM-D with active treatment, which amounts to a 54% response rate. However, in his third table on the HAM-D, he reports that 39 patients demonstrated the same 50% decrease with active treatment, which amounts to a 78% response rate. After going back to the Katz et al. (2011) re-analysis and then back to the original Katz et al. (2004) paper, it appears that there was a 62% response rate to DBI (apparently 16 of 26 patients) and a 46% response rate to paroxetine (11 of 24 patients); thus, the 54% overall response rate appears to be the correct one. As such, the third table on the HAM-D is probably presented incorrectly — the marginal probabilities appear to line up in a manner suggesting that rows and columns have been switched. If we transpose that third table according to this supposition, we arrive at the appropriate

54% response rate for active treatment. The numbers in the fourth 2x2 table that describes placebo also do not appear to add up correctly. The Katz et al. (2004) article indicates a 30% response rate for placebo (presumably 6 of 20 patients), yet the 2x2 table in Dr. Katz's response indicates that 15 of 19 patients recovered. Again, it appears that rows and columns were switched, and doing a transposition provides the reported results suggesting 6 patients responding to placebo over the course of the trial.

If my reorganization of these data is correct, we can now calculate the four NPP values needed to address Dr. Klein's question. These are as follows:

NPP=accuracy of a decision to discontinue medication at 2 weeks based upon < 20% improvement:

Active Drug:	NPP
Depressed mood-Retardation	88.2%
Anxiety-Agitation	74.5%
Hamilton Rating Scale	90.9%

Placebo:

Hamilton Rating Scale	75.0%
-----------------------	-------

What do these values tell us? It appears that in this limited sample, a lack of early response to Active Drug proves to have considerable negative predictive power, consistent with the viewpoint of Dr. Katz. It is reasonable to consider the HAM-D results as reflecting the most reliable and broad indicator of early response. If those non-early responders (on the HAM-D) had all been discontinued early in the trial, only 9.1% of these patients would have ultimately responded by the end of the trial. There are other utility considerations as to whether that "error rate" is acceptable, but the data do indicate that this early information about response on the HAM-D is quite predictive of eventual outcome. In addition, the 75% NPP for placebo response provides a reasonable rationale supporting a rapid discontinuation of placebo if there is no early response, a result that might be anticipated.

The limits of drawing broad conclusions about this issue from these data should be apparent. First, I am presuming that I have interpreted (or reinterpreted) the tables from Dr. Katz

correctly. Furthermore, the sample size is not sufficient to have confidence that these Bayesian estimates would be stable across other samples. In addition, this particular study was of two specific antidepressants. Because these Bayesian estimates are strongly influenced by a priori probabilities—here, the likelihood of a positive response to treatment over the course of the trial—these estimates would vary with different medications having different ultimate treatment efficacy, even if the early response profiles did not differ from those observed here. Even so, I believe this interchange is valuable as a demonstration of the need to consider the implication of treatment decisions at the level of the individual decision.

In sum, I believe that Dr. Klein is correct in suggesting that Dr. Katz et al. did not present the most informative analyses to answer the central question in his 2011 paper; that paper essentially presented sensitivity and specificity values, but what was needed to answer this question were positive predictive power and, more importantly for the discontinuation question, negative predictive power. Even so, when the proper figures are calculated, I believe that Dr. Katz is still correct — namely, that a decision to discontinue treatment following the lack of an early response is likely to be correct (based upon these limited data, and on the HAM-D) about 90% of the time.

December 17, 2015

Martin M. Katz's reply to Leslie C. Morey's comment

Dr. Morey's analysis uncovered an error in our last response, i.e., the figure of 39% of patients responding, was incorrect. The request for calculating "positive" and "negative percentage predictions" (NPP) is also, understandable. We have, however, as he agrees, already well established that the negative predictive value (NPP) is about 90%. That means that the finding that an absence of early improvement in the patient almost certainly leads to non-clinical response at outcome of treatment, is essentially validated. I and my co-authors have acknowledged in a previous paper that this is a small study and requires a larger study for confirmation of all the results. The Stassen et al. (1996) and Szegedi et al. (2009) studies that involved thousands of patients essentially confirmed the large part of those results. We do not believe that the statistics recommended by Morey will add significantly to what has already been learned, and to avoid

"overanalyzing" data from what was essentially a small study, are not inclined to conduct any further analyses.

References:

Stassen HH, Angst J, Delini-Stula A. Delayed onset of action of antidepressant drugs? Survey of recent results. *Eur Psychiatry*, 1996;12:166-76.

Szegedi A, Jansen WT, van Wugenburg AP. Early improvement in the first two weeks as predictors of treatment outcome in patients with major depressive disorder: a met-analysis including 6,562 patients. *J clin Psychiatry*, 2009;70:344-53.

January 14, 2016

Donald F. Klein's reply to Leslie C. Morey's comment

Dr. Morey's comment is somewhat difficult to follow. (Perhaps my trouble.) At any rate we agree that the tables presented by Dr. Katz present labeling problems. I disagree about the singular importance of NPV (NPP).

It is unfortunate that Dr. Morey did not relabel the 2 x 2 tables from which he derives his indices. With regard to the Hamilton scale my reconstruction is:

	$\leq 50\%$	$> 50\%$	
$\leq 20\%$	15	2	PPV = $25/(25+8) = 0.76$
$> 20\%$	8	25	NPV = $15/(15+2) = 0.88$
			Recovery rate $27/50 = 0.54$

Dr. Morey presents an NPP of 90.9% which seems close to my estimate of 0.88, but for such tables there should be no difference at all. Further, I can't follow the statement: "If those non-early responders (on the HAM-D) had all been discontinued early in the trial, only 9.1% of these patients would have ultimately responded by the end of the trial."

Dr. Morey seems to feel that a NPP approximating 90% justifies discontinuation early.

Certainly not in a clinical trial measured by ITT standards. Nor in the clinical case that often bears little resemblance to the study cases.

A close reading of Katz indicates that he was referring to a negative outcome across the range of measurements, rather than on a single measurement, to predict a bad outcome.

Finally, it should be noted that this is a single index and that no confidence limits are shown or suggested. Nor is it clear whether 95% limits or 90% limits, or broader, are in order.

Dr. Morey concludes: “I believe that Dr. Katz is still correct — namely, that a decision to discontinue treatment following the lack of an early response is likely to be correct (based upon these limited data, and on the HAM-D) about 90% of the time.”

In my last discussion of Marty’s statements, it is pointed out that the outcomes, based on his reported Hamilton scale measures can find no significant difference between “Active drug” and placebo. Therefore, this is a failed trial by FDA standards. Drawing any conclusion from it is most dubious.

March 31, 2016

*Martin M. Katz’s comment on Donald F. Klein’s reply to Leslie C. Morey’s
comment*

Drs. Klein and Morey have subjected our small study (Katz et al. 2011) to very careful analysis and disagree on the results. Klein questions Morey’s relying exclusively on the Hamilton Rating Scale to support the conclusion that 90% of patients who show no response to antidepressants within the first two weeks, will not respond to treatment at outcome. Morey has, therefore, supported the position that that drug treatment can be discontinued at the two-week point. Although I agree with Klein that further evidence from clinical methods other than the Hamilton Rating Scale, as we have applied them in these trials (e.g., Katz 2016), would be desirable before making such an important decision, we note that the 90% figure for the Hamilton has been replicated in several large sample studies involving thousands of patients. In view of the Hamilton method’s status as the established efficacy measure in the field, the evidence would

support Morey's and my conclusion that stopping the treatment at that two-week point is well justified.

Regarding the other points in Klein's comment, I have responded in a detailed manner previously. As for the issue of our combining the two drug groups, my co-investigators and I viewed that small step as justified for the 2011 analysis since that small study was a pilot. We, therefore, recommended that it be replicated in a prospective design with a more substantially sized patient sample. In view of this disagreement over the approach to analysis, we still recommend that route to resolving the issue.

More important, however, it is necessary to restate that the study they reanalyzed was on a relatively small patient sample and dealt with a very small piece of the overall problem. The main issue, the subject of this controversy, is whether the onset of clinical actions of efficacious antidepressant drugs lag several weeks beyond the almost immediate drug neurochemical effects in patients, or that those clinical actions begin to occur within the first two weeks of treatment. In reviewing the evidence, as pointed out previously, we turned to both independent studies (e.g., Coryell et al. 1982; Katz et al. 1987, 2004) which first reported significant early clinical actions and the later, more definitive large sample studies that included a range of antidepressants, such as those of Stassen et al. (1993) and Szegedi et al. (2009), for the answers. These methodologically sound studies provide uniform results. They establish that significant, clinically observable actions of the drugs when administered in the proper dosages in treatment-responsive patients are detected as early as the first week of treatment and on average within the first two weeks.

We believe this exchange on follow up analyses and views on these studies by several leaders in the field, most prominently by D. Klein, have served to further clarify the soundness of these results, and hopefully resolved the controversy on the main issue. The sooner these findings become accepted by the field at large, the sooner hundreds of patients already in severe pain, can be relieved of having to suffer through many weeks of ineffective treatments. The sooner, also, researchers can intensify their focus on the nature of these early, diverse clinical actions. This investigative approach should enhance our capacity to develop novel agents that hopefully, can act even more rapidly and be more broadly effective than the currently established antidepressant drugs.

References:

Coryell W, Coppen A, Zeigler VE, Briggs J. Early improvement as a predictor of response to amitriptyline and nortriptyline: a comparison of patient samples. *Psychological Medicine*, 1982;12:135-39.

Katz MM. *Clinical Trials of Antidepressants: How Changing the Model Can Uncover New, More Effective Molecules*. New York: Springer; 2016.

Katz MM, Berman N, Bowden CL, Frazer A. The componential approach enhances the effectiveness of 2-week trials for antidepressants. *J Clin Psychopharmacology*, 2011;31:253-4.

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

Katz MM, Tekell J, Bowden CL, Brannan S, Houston JP, Berman N, Frazer A. Onset and early behavioral effects of pharmacologically different antidepressants and placebo in depression. *Neuropsychopharmacology*, 2004;29:566-79.

Stassen HH, Angst J, Delini-Stula A. Time course of improvement under antidepressant treatment: a survival–analytic approach. *Eur Neuropsychopharmacol*. 1993;3:127-35.

Szegedi A, Jansen WT, van Wugenburg AP, van der Meulen E, Stassen HH, Thase ME. Early improvement in the first two weeks as predictors of treatment outcome in patients with major depressive disorder: a meta-analysis including 6,562 patients. *J Clin Psychiatry*, 2009;70:344-53.

June 9, 2016

Carlos Morra – Martin M. Katz – Donald F. Klein

Carlos Morra's comment on the interaction between Donald F. Klein and Martin M. Katz

I noted that Dr. Klein found that in his re-analysis of Dr. Katz's data the onset of antidepressant action was not established until the second or third week of treatment. The findings in his re-analysis differ also from Professor Kasper's, who found statistically significant change in the total score of the Hamilton Depression Scale during the first week (Kasper et al. 2006) and suggested that antidepressant action can be seen during the 24 hours after the first dose.

Reference:

Kasper S, Spadone C, Verpillat P, Angst J. Onset of action of escitalopram compared with other antidepressants: results of a pooled analysis. *Int Clin Psychopharmacol.* 2006;1:105-10.

May 28, 2015

Martin M. Katz's reply to Carlos Morra's comment

Professor Morra, in response to Klein's reanalysis of our data that suggested onset was generally delayed beyond two weeks, reports confirmatory findings of early response from Kasper et al. (2006). Their pooled, large sample study that used the Ham-D scale to evaluate change, added further to the accumulated evidence that clinical actions of the antidepressants can be detected within the first week of treatment. Their results also suggested that such clinical actions can be seen in many patients during the first 24 hours after the first dose.

It is useful to add the results of this Kasper study to the now firm body of evidence accumulated on this issue of time of onset of the clinical actions of antidepressants, vitally important to research on further drug development and to clinical practice in the treatment of the depressive disorders.

Reference:

Kasper S, Spadone C, Verpillat P, Angst J. Onset of action of escitalopram compared with other antidepressants: results of a pooled analysis. *Int Clin Psychopharmacol.* 2006;1:105-10.

November 19, 2015

Carlos Morra's response to Martin M. Katz's reply

Thank you for your reply. I would like to repeat that I don't know whether one can predict on the basis of your data within a week whether a particular depressed patient will respond to treatment with a particular antidepressant. Yet, I do know that many practicing psychiatrists share my opinion that from the clinical changes perceived, regardless of whether documented by rating scale scores, they can judge whether they should continue treatment after one week. I wonder about the justification for continuing treatment after one week without any encouraging clinical cues on

the basis of a questionable neuropharmacological theory with drugs, which may cause side effects in twice as many patients when prescribed than one can expect to favorably respond to them (Ban 2008).

Reference:

Ban TA. Comment on antidepressant medications and other treatments of depressive disorders: a CINP Task Force report based on a review of evidence. *The International Journal of Neuropsychopharmacology*, 2008;11:583-5.

August 11, 2016

Donald F. Klein's comment on Martin M. Katz's reply to Carlos Morra's comment

Dr. Morey and I both had trouble with Marty's response to my reply to Carlos Morra's comment in that the tables were insufficiently labeled. The section following is a copy of the tables and Marty's description.

Below are the 2X2 Tables for the active drugs, the placebo and the chi-square results. The rows are number of "early improvements" (>20%), the columns are number of recovered (>50%) at outcome.

I. Active Drugs

Depressed mood-retardation		
Early Improvement (EI)	Recovered (>50% Ham-D decrease)	
No	Yes	
15	2	17 chi-square =18.5
8	25	33 p<0.0001
23	27	50

Using this as an example, if we transcribe and label which row represents >20%. Say it is Row 1. Since we know that there will be a positive correlation between greater than 20% and

greater than 50% this infers that column 1 is greater than 50%.

	C1	C2
Row 1	15	2
Row 2	8	25

PPV is Positive Predictive Value, i.e., of those predicted to be positive, the proportion that is actually positive.

	>50%	<= 50%	
>20%	15	2	PPV= 15/ (15+2) =.88
<=20%	8	25	

Reasonable - but 17 were predicted to do well and 27 did-a marked under prediction however if we had guessed. Looks that Row 1 is <=20% it would be sensible to reverse the column labels to preserve the predicted positive correlation

<=50%	> 50%	
<=20%	15	2
>20%	8	25 Recovery rate 27/50=.54

The difference in PPVs is enough to indicate that Marty's directions are ambiguous, as Dr. Morey also found. Note 33 are predicted to do well but only 27 did, an over prediction. This approximates Marty's statement that, "It was further demonstrated in these studies that this amount of 'early improvement', i.e., >20% decrease, was clinically significant in that it could predict that 70% of patients showing this early improvement would go on to respond at 6 or 8 weeks to the experimental treatment." So, the second table is probably the correct one. However, it is unclear to me how this over-prediction means that it is clinically significant. Further, there is no contrast with placebo either in the: paper by Katz MM, Berman N, Bowden CL, Frazer A., or here.

II. Placebo Treatment Group

	Hamilton Rating Scale		
	No	Yes	
	EI Recovered		
3	10	13	chi-square =0.102
1	5	6	p<0.75
4	15	19	

Taken literally this seems to indicate that 15/19 recovered on placebo. Dr. Morey puzzles over this: “The Katz et al. (2004) article indicates a 30% response rate for placebo (presumably 6 of 20 patients), yet the 2x2 table in Dr. Katz’s response indicates that 15 of 19 patients recovered. Again, it appears that rows and columns were switched, and doing a transposition provides the reported results suggesting 6 patients responding to placebo over the course of the trial.” This seems reasonable. Col 2 should be 1, 5 to preserve the chi-square the Table looks like. So, 6/19 recovered

	<=50%	> 50%	
<=20%	10	5	Recovery rate 6/19= .32
>20%	3	1	

	Drug	Placebo
Rec	27	6
Nrec	25	13

Contrasting the two recovery rates available, chi-square = 2.3, which is far from significant and casts doubt on any “finding.”

Other major problems: this table is referred to as active drug. The results somehow combine the 24 in Paroxetine study with the 26 in DMI study. No justification is given for this. Since Paroxetine was picked as a serotonergic agent and DMI as a Noradrenergic agent, the combination is really strange. This table is not what we asked for which was the individual studies.

I thank Marty for providing the placebo data as used in the calculations by Dr. Morey. That this non-significant 6-week contrast is held to justify a much shorter trial escapes me.

January 21, 2016

Martin M. Katz's reply to Donald F. Klein's comment on Katz's reply to Carlos Morra's comment

Dr. Klein has gone to great pains to reanalyze the data from our paper (Katz et al. 2011) that reported on a secondary analysis from the 2004 Texas study. I and colleagues consider that we did well by these results, as previously explained, but understand that others can have different perspectives on how to go about it. Nevertheless, as explained in my last note, we stand by the reported results, decided that the study has had sufficient scrutiny, and we would avoid any further “over-analysis” of this relatively small size sample study.

More important, however, this dispute about how to analyze this small study, serves to distract from the main thrust of the controversy over “the onset of antidepressant clinical actions.” In my initial essay on the topic, I summarized the evidence which stems not from this small study, but from many studies, including at least two meta-analyses that involved thousands of patients, the entire range of antidepressant drugs and the necessary placebo controls, i.e., Stassen et al. 1993, Szegedi et al. 2009. Among these studies was the Texas study (Katz et al. 2004) that had as its major aim the determination of onset. That study also contained a placebo control and met all requirements of a controlled, direct study of the issue. The results of these many studies are summarized in my essay with the appropriate references. Included also in that review is the chronology of events and the studies that resulted in the controversy. In brief, that initial summary began with early clinical observations by such astute clinical investigators as Kuhn, in his original paper (1958) and Paul Kielholz (1968), who all observed positive clinical effects by the end of the first week of treatment; the Quitkin et al. 1984 study, not initially designed to determine “onset,” found it would take several weeks for drug-induced clinical effects to be detected; through the Katz et al. 1987 study, which although also not designed to identify onset, reported significant clinical actions to occur within 1 to 2 weeks. The latter study did not have a placebo group and was not designed to determine onset, so the results were not

apparently acceptable to Dr. Klein, one of the coauthors of the Quitkin et al. study. In the follow-up studies, starting with the Stassen et al. study in 1993, and then later, Szegedi et al. tackle the problem of onset directly. All of these studies came up with roughly the same results, i.e., onset occurs within the first two weeks of treatment. The work of Szegedi and Stassen among others cited also found that “failure of the patient to respond with early improvement (within 2 weeks) almost certainly (>90%) results in failure of clinical response at treatment outcome.”

Responding to a published critique of our work by the Quitkin, Klein et al. group in the early 90s, we did an analysis of their 1984 study that had reported a “lag” in onset of clinical actions, delayed for several weeks, the work that appeared to ignite the controversy. We pointed out (Katz et al. 1997) that the study, allowing for the fact that it was not targeted on the onset issue, was flawed in achieving that goal in several respects, notably (1) study dosage of drugs was relatively low, administered gradually during the first two weeks, so that the effective dosage was not reached until 2 weeks; (2) the study outpatients were only mild to moderately depressed; (3) measures of clinical actions during the first two weeks were inadequate. Quitkin, Klein and colleagues apparently disturbed by this analysis published a not very convincing rejoinder. (See exchange in *Neuropsychopharmacology* referenced in the initial essay).

I have returned the discussion to the central theme of this controversy because the over-scrutinizing of the secondary analysis from our 2004 study has clouded the basic issues, and diverted attention from the main, indisputable facts on the determination of onset, supported by the several meta-analytic and directly targeted studies on the issue. I am now satisfied that I have presented the case as clearly as possible.

It is most important to get the facts straight on this issue since the “time of onset” is critical, not only for clinicians for effective clinical practice, but for basic investigators attempting to uncover the mechanisms underlying therapeutic drug actions and seeking to develop new more effective drugs.

References:

Katz MM, Berman N, Bowden CL, Frazer A. The componential approach enhances the effectiveness of 2-week trials of new antidepressants. *J Clin Psychopharmacology*, 2011;37:193-218.

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

Katz MM, Tekell J, Bowden CL, Brannan S, Houston JP, Berman N, Frazer A. Onset and early behavioral effects of pharmacologically different antidepressants and placebo in depression. *Neuropsychopharmacology*, 2004;29:566-79.

Katz MM, Bowden CL, Stokes P, Casper R, Frazer A, Koslow S, Kocsis J, Secunda S, Swann A, Berman N. Can the effects of antidepressants be observed in the first two weeks of treatment? *Neuropsychopharmacology*, 1997;17:110-1.

Kuhn R. The treatment of depressive states with G22355 (imipramine hydrochloride). *American Journal of Psychiatry*, 1958;115:459-64.

Kielholz P. Die Behandlung endogener Depressionen mit Psychopharmaka. *Deut Med Wschr.* 1968;93:701.

Quitkin F M, Rabkin JG, Ross D, Stewart JW. Identification of true drug response to antidepressant. *Arch Gen Psychiatry*, 1984;41:782-6.

Redmond E. The timing, specificity, and clinical prediction of tricyclic drug effects in depression. *Psychol Med.* 1987;17:297-309.

Stassen HH, Angst J, Delini-Stula A. Time course of improvement under antidepressant treatment: a survival-analytic approach. *Eur Neuropsychopharmacol.* 1993;3:127-35.

Szegedi A, Jansen WT, van Wugenburg AP, van der Meulen E, Stassen HH, Thase ME. Early improvement in the first two weeks as predictors of treatment outcome in patients with major depressive disorder: a meta-analysis including 6,562 patients. *J Clin Psychiatry*, 2009;70:344-53.

March 17, 2016

Donald F. Klein's corrected comment on Martin M. Katz's reply to Carlos Morra's comment

Dr. Morey and I both had trouble with Marty's response to my reply to Carlos Morra's comment in that the tables were insufficiently labeled. The section following is a copy of the tables and Marty's description.

“Below are the 2X2 Tables for the active drugs, the placebo and the chi-square results. The rows are number of ‘early improvements’ (>20%), the columns are number of recovered

(>50%) at outcome.”

I mistakenly thought that analysis of “Depressed mood-retardation” would be to the point. I did not realize that analysis of Hamilton score would be better in terms of comparability with placebo group. Below the Hamilton scale is used. The first two lines supplied by Marty immediately above the table were deleted as simply confusing.

I. Active Drugs

Assuming that row one is the predication of greater than 20%

Hamilton Rating Scale

	>50%	≤ 50%	
>20%	15	2	PPV= $15/(15+2) = 0.88$
≤20%	8	25	NPV= $25/(25+8) = 0.758$ Recovery rate = 0.46

17 were predicted to do well but 27 did - a marked under prediction.

Conversely, if Row 1 is ≤20%, the column labels must be reversed to preserve the positive correlation

	≤50%	> 50%	
≤20%	15	2	PPV= $25/(25+8) = 0.76$
>20%	8	25	NPV = $15/(15+2) = 0.88$ Recovery rate $27/50 = 0.54$

The difference in PPV and NPV Recovery Rate are enough to indicate that Marty's directions are ambiguous, as Dr. Morey also found.

Note that 33 are predicted to do well but only 27 did, an over prediction. $27/33 = 0.82$. This approximates Marty's statement that, "It was further demonstrated in these studies that this amount of 'early improvement,' i.e., >20% increase, was clinically significant in that it could predict that 70% of patients showing this early improvement would go on to respond at 6 or 8 weeks to the experimental treatment."

So, the second table is probably the correct one. However, it is unclear to me how this over-prediction means that it is clinically significant. Further, there is no contrast with placebo either in the paper by Katz MM, Berman N, Bowden CL, Frazer A. or here.

II. Placebo Treatment Group

	Hamilton Rating Scale		
	EI Recovered		
No	Yes		
3	10	13	chi-square =0.102
1	5	6	p<0.75
4	15	19	

Taken literally, this seems to indicate that 15/19 recovered on placebo. Dr. Morey puzzles over this: "The Katz et al. (2004) article indicates a 30% response rate for placebo (presumably 6 of 20 patients), yet the 2x2 table in Dr. Katz's response indicates that 15 of 19 patients recovered. Again, it appears that rows and columns were switched, and doing a transposition provides the reported results, suggesting 6 patients responding to placebo over the course of the trial." This seems reasonable - Col 2 should be 1 and 5. To preserve the chi-square the Table looks like: So, 6/19 recovered.

	$\leq 50\%$	$> 50\%$	
20%	10	5	Recovery rate $6/19 = 0.32$
>20%	3	1	

Contrasting the recovery rates of active drug and placebo we find:

Drug	Placebo
------	---------

Recover	27	6	33
Nrec	23	13	36
	50	19	

Contrasting the two recovery rates, chi-square = 2.77, far from significant. This casts doubt on any “finding” that Marty proposes.

Other major problems remain. This table is referred to as active drug = 50. This combines the 24 in the Paroxetine study with the 26 in the DMI study. No justification is given for this. Since Paroxetine was picked as a serotonergic agent and DMI as a noradrenergic agent, the combination is really strange. This table is not what we asked for, which was the individual studies.

I thank Marty for providing the placebo data as used in the calculations by Dr. Morey. That this non-significant 6-week contrast is held to justify a much shorter trial escapes me.

March 24, 2016

Martin M. Katz's reply to Donald F. Klein's corrected comment on Katz's reply to Morra's comment

Morra found results of Zarate et al. (2006) study, which showed ketamine to effect rapid onset of antidepressant actions in some patients within the first day of treatment, to support soundness of earlier finding that response during first two weeks was predictive of a positive outcome in trials of putative antidepressants. Klein is right to question whether the Zarate study any provides evidence in support of this predictive hypothesis. Zarate et al.'s conclusion (2006) refers to an onset achieved through a single intravenous dose of ketamine within two hours, “the effects remaining significant for one week”, and does not deal, as Klein points out, with the outcome issue.

Morra's earlier comment about Kasper et al.'s (2006) evidence regarding onset within the first week, however, supports one of the basic issues in this controversy, i.e., that clinical onset in treatment-responsive patients occurs early within the first two weeks of treatment.

The bottom line here, which Klein still seems to reluctant to accept, despite strong accumulating evidence from studies that investigated multiple antidepressants across thousands of

soundly, diagnosed depressive disorders (e.g., Stassen et al. 1997, Szegedi et al. 2009, Katz et al. 2004), is that early reactivity or non-reactivity to the drugs, i.e., within two weeks of treatment, will predict at a high level of confidence, which patients will respond to the treatment at outcome of a 6-12-week trial.

References:

Kasper S, Spadone C, Verpillat P, Angst J. Onset of action of escitalopram compared with other antidepressants: results of a pooled analysis. *Int Clin Psychopharmacol.* 2006;1:105-10.

Katz MM, Tekell J, Bowden CL, Brannan S, Houston JP, Berman N, Frazer A. Onset and early behavioral effects of pharmacologically different antidepressants and placebo in depression. *Neuropsychopharmacology*, 2004;29:566-79.

Stassen HH, Angst J, Delini-Stula A. Delayed onset of action of antidepressant drugs? Survey of recent results. *Eur Psychiatry*, 1997;12:166-76.

Szegedi A, Jansen WT, van Wugenburg AP, van der Meulen E, Stassen HH, Thase ME. Early improvement in the first two weeks as predictors of treatment outcome in patients with major depressive disorder: a meta-analysis including 6,562 patients. *J Clin Psychiatry.* 2009;70(3):344-53.

Zarate CA Jr, Singh JB, Carlson PJ, Brutsche NE, Ameli R, Luckenbaugh DA, Charney DS, Manji HK. A randomized trial of an N-methyl-D-aspartate antagonist in treatment-resistant major depression. *Arch Gen Psychiatry*, 2006; 63:856-64.

July 14, 2016

Donald F. Klein reply to Martin M. Katz's comment on Klein's reply to Morra's comment

Our previous discussions were about drugs whose clinical actions seem somehow related to the fast blockade of reuptake mechanisms for norepinephrine and serotonin. The interesting fast effects of ketamine are still poorly understood. Comparing the various aspects of the ketamine response to our older antidepressants seems premature

However, Katz and I agree that the Zarate et al (2006) study, which showed ketamine's rapid onset of antidepressant actions in some patients was irrelevant to Katz' claim that response during the first two weeks was predictive of a positive outcome in trials of putative antidepressants.

Katz argues, "Morra's earlier comment about Kasper et al's (2006) evidence regarding onset

within the first week, however, supports one of the basic issues in this controversy, i.e., that clinical onset in treatment-responsive patients occurs early within the first two weeks of treatment.”

This is difficult to follow. Kasper et al.’s abstract begins, “In general, antidepressant drugs are regarded as too slow acting. Most patients who benefit from treatment require more than 2 weeks of therapy to respond to treatment. An efficacious and well-tolerated antidepressant drug with an earlier onset of effect would be of greater interest to clinicians and patients.”

Clearly Kasper does not agree with Katz. His study was about whether escitalopram’s effect (undefined) had a faster onset than other SSRIs. The key finding was, “The mean change in MADRS total scores was significantly higher for escitalopram-treated patients than for patients treated with the comparators on day 7 (-3.9 versus -3.4, respectively, $P = 0.029$).” This indicates a difference (0.5), implicitly considered by Katz as a measure of depression alleviation occurring within the first week, thus supporting his claim.

However, is that accurate? Item analysis revealed “that ... inner tension was significantly decreased” in the escitalopram group compared to the all comparators group as early as week 1 ... also at week 2.

Reduced sleep was similarly affected as early as week 2 ... This difference remained significant until the end of the trial.”

Are these clear indicators of an antidepressant effect? An alternative conclusion is that escitalopram is on average, a better faster acting sedative than the other SSRIs. However, without a placebo contrast even this is obscure. Perhaps the pooled SSRIs had an anti-sedation effect, muffled by a positive placebo effect and escitalopram was simply a placebo.

In any case, strong support for Katz’s hypothesis of antidepressant effect occurring within the first two weeks receive no support. The other hypothesis that early improvement was predictive of positive outcome was not even alluded to. It is difficult to understand why this article is considered supportive of Katz’s views.

Marty sees me as “reluctant to accept, despite strong accumulating evidence... (e.g., Stassen et al 1997, Szegedi et al 2009, Katz et al 2004), ... that early reactivity or non-reactivity to the drugs, i.e., within two weeks of treatment, will predict at a high level of confidence, which patients will respond to the treatment at outcome of a 6 to 12 week trial.”

Right, reluctance puts it mildly. In my comments every single one of the articles that Katz cites as supportive were found irrelevant or non-supportive. This 4 year discussion might be shortened by direct

confrontations over the specific data and analyses that Katz and I disagree about. The grounds for reluctance are not obscure.

References:

Kasper S, Spadone C, Verpillat P, Angst J. Onset of action of escitalopram compared with other antidepressants: results of a pooled analysis. *Int Clin Psychopharmacol*. 2006; 1:105-10.

Katz MM, Tekell J, Bowden CL, Brannan S, Houston JP, Berman N, Frazer A. Onset and early behavioral effects of pharmacologically different antidepressants and placebo in depression. *Neuropsychopharmacology* 2004; 29: 566-79.

Stassen HH, Angst J, Delini-Stula A. Delayed onset of action of antidepressant drugs? Survey of recent results. *Eur Psychiatry* 1997; 12: 166-76.

Szegedi A, Jansen WT, van Wugenburg AP. Early improvement in the first two weeks as predictors of treatment outcome in patients with major depressive disorder: a meta-analysis including 6,562 patients. *J Clin Psychiatry* 2009; 70: 344-53.

Zarate CA Jr, Singh JB, Carlson PJ, Brutsche NE, Ameli R, Luckenbaugh DA, Charney DS, Manji HK. A randomized trial of an N-methyl-D-aspartate antagonist in treatment-resistant major depression. *Arch Gen Psychiatry* 2006; 63: 856-64.

September 8, 2016

Carlos Morra - Donald F. Klein – Martin M. Katz

Donald F. Klein's reply to Carlos Morra's comment

Dr. Morra believes that our stand regarding antidepressant onset rules out evidence of first week effect. Not so, our general point is, even if first week anti-depressant effect is actually detected (given the ample opportunity for misleading early drug vs placebo effects in the first week that have no relation to later clinically documented effects at 6 weeks), the central issue is that most substantial onsets of antidepressant effects occur later than one week. Is that a factual issue? I think no one has denied that.

The heuristic importance is that we agree that the drug effect on reuptake blockade is immediate, causing greater synaptic neurotransmitter concentration. The issue is whether this is sufficient for clinical benefit or is at best the first domino. If sufficient, the quest for better anti-

depressants should hover about reuptake blockade, synaptic concentration, immediate receptor stimulation, etc. In fact, this is the accepted model for drug development, but clinical outcomes over forty years have not improved and no new class of antidepressant has been discovered...

If the drug's immediate synaptic effects are usually at some distance from anti-depressant onset then the simple model fails. It also fails since anti-depressants with the same immediate synaptic effects do not make normal subjects happier. Evidently the drug is in some fashion neutralizing the pathophysiology and clinical benefit depending on the extent of the neutralization.

Unfortunately, models that incorporate the drug effect over time with the still unknown pathophysiology of the manifest illness, (my guess is an adaptive hedonic mechanism that requires remedy of damaged stabilizing negative feedback loops), have not testably replaced the simplistic model. I attempted from Marty Katz's report to elicit, by reconstruction, the data relating the week two effect to the six-week outcome. Marty says my algebra was cockeyed and he may well be correct. Fortunately, this is beside the point. My data reconstruction may well be wrong. This lacks importance when the real *actual data is in Marty's data bank. My substantial point is that if Marty revealed the actual data in 2 X 2 form, relating two-week effects to six- week outcome, the discrepancy from his theories would be glaring. Also, it would make evident that consideration of the placebo effect was missing. Fortunately, Marty has the opportunity to demonstrate the correctness, or not, of his theories by this simple data demonstration. Let's see it*

July 23, 2015

Carlos Morra's response to Donald F. Klein's reply to Morra's comment

My intention with my comment was to bring to attention that current findings in this area of research are in favor of Marty Katz's contention that one can detect antidepressant effect after one week of treatment, or even sooner. Whether one can do it on the basis of his data, I don't know.

I agree with Dr. Klein that the hypothesis or belief that a drug-induced increase of a neurotransmitter in the synaptic gap is responsible for antidepressant effects has not paid off in terms of drug development. But, then the hypothesis or belief that postsynaptic changes ("cascade") are responsible for antidepressant effects, has not paid off either. In recent years,

ketamine administration, allegedly by its effect on ion channels, resulted in improvement of depression within 110 minutes (Maeng and Zarate 2007; Zarate et al. 2003, 2006).

Nothing of this has, of course, anything to do with whether one can predict antidepressant effects from Dr. Katz's data after one week of treatment and that is at the heart of this exchange. But, to repeat: my intention with my comment was to bring to attention that some current findings are in favor of the contention that, in general, one can predict whether a patient will respond to antidepressant treatment within the time frame Marty Katz suggests, even if Dr. Klein's argument is correct that the substantial response to treatment takes place later.

References:

Maeng S, Zarate CA Jr. The role of glutamate in mood disorders: results from the ketamine in major depression study and the presumed cellular mechanism underlying its antidepressant effects. *Curr Psychiatry Rep.* 2007;9:467-74.

Zarate CA Jr, Du J, Quiroz J, Gray NA, Denicoff KD, Singh J, Charney DS, Manji HK. Regulation of cellular plasticity cascades in the pathophysiology and treatment of mood disorders: role of the glutamatergic system. *Ann N Y Acad Sci.* 2003;1003:273-91.

Zarate CA Jr, Singh JB, Carlson PJ, Brutsche NE, Ameli R, Luckenbaugh DA, Charney DS, Manji HK. A randomized trial of an N-methyl-D-aspartate antagonist in treatment-resistant major depression. *Arch Gen Psychiatry.* 2006; 63:856-64.

April 14, 2016

Donald F. Klein's response to Carlos Morra's response to Klein's reply to Morra's comment

You state: "one can predict whether a patient will respond to antidepressant treatment within the time frame Marty Katz suggests, even if Dr. Klein's argument is correct that the substantial response to treatment takes place later" and cite three articles in support.

Maeng and Zarate (2007) appear to agree with me about the irrelevance of week one findings in the sort of antidepressants Katz had studied, by stating: "Although most antidepressants exert their initial effects by increasing the intrasynaptic levels of serotonin and/or norepinephrine, the resolution of core depressive symptoms becomes manifest only after weeks of

administration." They hypothesize the acute and maintained effect of ketamine infusion, "are due to an increase in AMPA relative to NMDA glutamatergic throughput, which results in increased synaptic potentiation." In their cited Berman pilot study, using their 50% HAMD decrement criterion, 4/7 subjects had this immediate response. There is no prediction of this benefit or of any degree of maintained benefit.

Zarate et al. (2003) simply suggest that glutamatergic mechanisms may be of anti-depressant relevance.

Zarate et al. (2006) describe a crossover study of IV ketamine vs placebo, two weeks apart. Using the 50% fall in HDRS as criterion, about 55% of the ketamine treated group achieved this during the first post-infusion day.

Again, there was no outcome predictability evident.

Of interest, although not strictly relevant to this discussion, "The earliest improvements were at 40 minutes for depressed mood and guilt. Depersonalization or derealization was worse from 40 to 110 minutes. Motor retardation and gastrointestinal symptoms were worse at 40 minutes, but at day 1, motor retardation was better for participants receiving ketamine than those receiving placebo."

This review does not agree with your statement, "one can predict whether patient will respond to antidepressant treatment within the time frame Marty Katz suggests." That ketamine is having quick acute benefits is not to the point. That neither these acute or maintained benefits, due to ketamine, have been predictable is interesting but also irrelevant to the discussion that Katz and I had.

References:

Zarate CA Jr, Du J, Quiroz J, Gray NA, Denicoff KD, Singh J, Charney DS, Manji HK. Regulation of cellular plasticity cascades in the pathophysiology and treatment of mood disorders: role of the glutamatergic system. *Ann N Y Acad Sci.* 2003;1003:273-91.

Zarate CA Jr, Singh JB, Carlson PJ, Brutsche NE, Ameli R, Luckenbaugh DA, Charney DS, Manji HK. A randomized trial of an N-methyl-D-aspartate antagonist in treatment-resistant major depression. *Arch Gen Psychiatry.* 2006;63:856-64.

July 7, 2016

Carlos Morra's response 2 to Donald F. Klein's response to Carlos Morra's response to Klein's reply to Morra's comment

Thank you for clarifying the findings of Zarate and his associates and some findings on which the suggested antidepressant effect of ketamine is based. I am wondering whether you would be willing to elaborate further on your thoughts about the findings of Kasper and his associates who reported on statistically significant change in the total score of the Hamilton Depression Scale during the first week of treatment and suggested that antidepressant action can be seen during the 24 hours after the administration of the first dose?

References:

Kasper S, Spadone C, Verpillat P, Angst J. Onset of action of escitalopram compared with other antidepressants: results of a pooled analysis. *Int Clin Psychopharmacol*. 2006;1:105-10.

Zarate CA Jr, Du J, Quiroz J, Gray NA, Denicoff KD, Singh J, Charney DS, Manji HK. Regulation of cellular plasticity cascades in the pathophysiology and treatment of mood disorders: role of the glutamatergic system. *Ann N Y Acad Sci*. 2003;1003:273-91.

Zarate CA Jr, Singh JB, Carlson PJ, Brutsche NE, Ameli R, Luckenbaugh DA, Charney DS, Manji HK. A randomized trial of an N-methyl-D-aspartate antagonist in treatment-resistant major depression. *Arch Gen Psychiatry*, 2006; 63:856-64.

July 21, 2016

Martin M. Katz's comment on Donald F. Klein's reply to Carlos Morra's comment

Klein requested that we show the actual data from a table in our (Katz et al. 2011) paper that showed "early improvement" (EI) at two weeks of treatment to be predictive of treatment outcome after six weeks of antidepressant treatment. This data showing the two-week improvement and 6-week outcome ratings in a 2X2 table, could then be used to test directly the significance of the relationship between early improvement and outcome, i.e., to confirm or disconfirm our prediction about the predictability of early treatment response.

We have since tested the predictability of changes, i.e., $\geq 20\%$, at two weeks for the two antidepressant drugs, desipramine and paroxetine ($n=50$), in the 2004 study (Katz et al.). We defined recovery at outcome as $\geq 50\%$ decreases in each of two severity dimensions from the MV method (Katz et al. 2004) and the Ham-D (21 item) total score. The two dimensions were “depressed mood-retardation” and “anxiety-agitation-somatization”. A comparable analysis was run for the group treated with placebo on the Ham-D.

Below are the 2X2 Tables for the active drugs, the placebo and the chi-square results. The rows are number of “early improvements” ($\geq 20\%$), the columns, number of recovered ($\geq 50\%$) at outcome.

I. Active Drugs

Depressed mood-retardation

Early Improvement(EI)	Recovered ($>50\%$ Ham-D decrease)		
	No	Yes	
15	2	17	chi-square =18.5
8	25	33	p<0.0001
23	27	50	

Anxiety-Agitation

17	7	24	chi-square =11.5
6	20	26	p<0.0007
23	27	50	

Hamilton Rating Scale

10	13	23	chi-square =11.5
1	26	27	p<0.0007
11	39	50	

II. Placebo Treatment Group

Hamilton Rating Scale

EI	Recovered		
	No	Yes	
3	10	13	chi-square =0.102
1	5	6	p<0.75
4	15	19	

Results: We note that all chi-square tests for the active drugs were significant well beyond 0.001 probability level affirming the strong relationship reported earlier in the paper describing the 2004 study (Katz et al.), They, therefore, do not negate as Klein expected, but further support the prediction hypothesis tested earlier in our study.

References:

Katz MM, Tekell JL, Bowden CL, Brannan S, Houston JP, Berman N, Frazer A. Onset and early behavioral effects of pharmacologically different antidepressants and placebo in depression. *Neuropsychopharmacology*, 2004;29:566-79.

Katz MM, Houston JP, Brannan S, Bowden CL, Berman N, Swann A, Frazer A. A multivantaged behavioral method for measuring onset and sequence of the clinical actions of antidepressants. *International Journal of Neuropsychopharmacology*, 2004;7:471-9

Katz MM, Berman N, Bowden CL, Frazer A. The componential approach enhances the effectiveness of 2-week trials of new antidepressants. *J Clin Psychopharmacology*, 2011;37:193-218.

October 15, 2015

Donald F. Klein's reply to Martin M. Katz's comment on Klein's reply to Morra's comment

Our previous discussions were about drugs whose clinical actions seem somehow related to the fast blockade of reuptake mechanisms for norepinephrine and serotonin. The interesting fast effects of ketamine are still poorly understood. Comparing the various aspects of the ketamine response to our older antidepressants seems premature

However, Katz and I agree that the Zarate et al. (2006) study, which showed ketamine's rapid

onset of antidepressant actions in some patients was irrelevant to Katz' claim that response during the first two weeks was predictive of a positive outcome in trials of putative antidepressants.

Katz argues, "Morra's earlier comment about Kasper et al.'s (2006) evidence regarding onset within the first week, however, supports one of the basic issues in this controversy, i.e., that clinical onset in treatment-responsive patients occurs early within the first two weeks of treatment."

This is difficult to follow. Kasper et al.'s abstract begins: "In general, antidepressant drugs are regarded as too slow acting. Most patients who benefit from treatment require more than 2 weeks of therapy to respond to treatment. An efficacious and well-tolerated antidepressant drug with an earlier onset of effect would be of greater interest to clinicians and patients."

Clearly Kasper does not agree with Katz. His study was about whether escitalopram's effect (undefined) had a faster onset than other SSRIs. The key finding was: "The mean change in MADRS total scores was significantly higher for escitalopram-treated patients than for patients treated with the comparators on day 7 (-3.9 versus -3.4, respectively, $P = 0.029$)." This indicates a difference (0.5), implicitly considered by Katz as a measure of depression alleviation occurring within the first week, thus supporting his claim.

However, is that accurate? Item analysis revealed "that... inner tension was significantly decreased" in the escitalopram group compared to the all comparators group as early as week 1... also at week 2.

Reduced sleep was similarly affected as early as week 2. This difference remained significant until the end of the trial.

Are these clear indicators of an antidepressant effect? An alternative conclusion is that escitalopram is on average, a better faster acting sedative than the other SSRIs. However, without a placebo contrast even this is obscure. Perhaps the pooled SSRIs had an anti-sedation effect, muffled by a positive placebo effect and escitalopram was simply a placebo.

In any case, strong support for Katz's hypothesis of antidepressant effect occurring within the first two weeks receive no support. The other hypothesis that early improvement was predictive of positive outcome was not even alluded to. It is difficult to understand why this article is considered supportive of Katz's views.

Marty sees me as "reluctant to accept, despite strong accumulating evidence... (e.g., Stassen et al. 1997, Szegedi et al. 2009, Katz et al. 2004), ... that early reactivity or non-reactivity to the drugs, i.e., within two weeks of treatment, will predict at a high level of confidence, which patients will respond to

the treatment at outcome of a 6 to 12 week trial.”

Right, reluctance puts it mildly. In my comments every single one of the articles that Katz cites as supportive were found irrelevant or non-supportive. This 4-year discussion might be shortened by direct confrontations over the specific data and analyses that Katz and I disagree about. The grounds for reluctance are not obscure.

References:

Kasper S, Spadone C, Verpillat P, Angst J. Onset of action of escitalopram compared with other antidepressants: results of a pooled analysis. *Int Clin Psychopharmacol*. 2006;1:105-10.

Katz MM, Tekell J, Bowden CL, Brannan S, Houston JP, Berman N, Frazer A. Onset and early behavioral effects of pharmacologically different antidepressants and placebo in depression. *Neuropsychopharmacology*, 2004;29:566-79.

Stassen HH, Angst J, Delini-Stula A. Delayed onset of action of antidepressant drugs? Survey of recent results. *Eur Psychiatry*, 1997;12:166-76.

Szegedi A, Jansen WT, van Wugenburg AP. Early improvement in the first two weeks as predictors of treatment outcome in patients with major depressive disorder: a meta-analysis including 6,562 patients. *J Clin Psychiatry*, 2009;70:344-53.

Zarate CA Jr, Singh JB, Carlson PJ, Brutsche NE, Ameli R, Luckenbaugh DA, Charney DS, Manji HK. A randomized trial of an N-methyl-D-aspartate antagonist in treatment-resistant major depression. *Arch Gen Psychiatry*, 2006; 63:856-64.

September 8, 2016

Martin M. Katz's comment on Carlos Morra's response to Donald F. Klein's reply to Morra's comment

Carlos Morra agrees that reported findings from several well-controlled studies firmly support significant clinical actions of antidepressant drugs as occurring within the first two weeks of treatment. He agrees with D. Klein that the hypothesis linking the drug effective actions to inhibition of reuptake of neurotransmitters at synapses, is dubious, thus, disappointing in its implications for new drug development. He cites findings, however, with ketamine of rapid antidepressant action that indicate the drug effect probably due to another cellular mechanism.

This does not invalidate the main issue of whether clinical response at outcome can be predicted from patient reactivity to treatment during the first two weeks. In essence, he supports the main thesis of the onset controversy. Thus, he takes seriously, as we do, that adopting this finding that outcome can be predicted from two-week's response, should lead to shorter trials, and prevent long suffering patients from being exposed to unnecessary, additional weeks of ineffective treatments. Adopting this procedure will, as in the case of ketamine, help open pathways to the discovery of new, more effective and more rapidly acting drugs.

May 26, 2016

Elemer Szabadi – Martin M. Katz – Donald F. Klein

Elemer Szabadi's comment

In the debate between Martin Katz and Donald Klein, the question was raised whether the finding of an antidepressant effect during the first couple of weeks of a clinical trial of an antidepressant “could be applied to individual patients” (see Katz, INHN February 26, 2015). Some observations made by us a few years ago (Szabadi et al. 1976) may be relevant to this question. We measured pause time (total time taken up by the nine inter-phonation pauses in a sample of automatic speech, i.e. counting from one to ten) in depressed patients both before, during and after treatment with an antidepressant (amitriptyline); speech measures (both phonation and pause times) and Hamilton Depression Scores were recorded at weekly intervals. There was a close correlation between pause time and depression scores: as the depression subsided, the counting speeded up, until, on recovery from depression, it stabilized at a level characteristic of the patient. Although initially we proposed pause time as a measure of psychomotor retardation, later, due to its close coupling to mood, we suggested that, in fact, it may be “an objective sensitive correlate of mood” (Szabadi and Bradshaw 1983). In our early study we already found, albeit in a small number of individual patients and without a placebo control, an appreciable reduction in both Hamilton scores and pause time, during the first week of treatment. In fact, the scores declined monotonically until full remission was observed after about six weeks. The usefulness of speech pause time as an index of psychomotor retardation was confirmed by others (Greden and Carroll 1980; Greden et al. 1981; Hoffmann et al. 1985; for a comprehensive review, see Bennabi et al.

2013.) It is of interest that other authors have also found reductions in this measure early in the course of antidepressant treatment. For example, Greden and Carroll (1980) commented: “of potential clinical importance, in two cases we documented decreases in pause time of more than 50% within several days after starting treatment.” Interestingly, a similar degree of reduction in depression and pause time scores was observed in the patient whose data we published in a figure (Szabadi et al. 1976). However, not every patient may show such dramatic improvement early on in the course of antidepressant treatment, and, although improvement may start soon after initiation of treatment, it may not be large enough to be of clinical significance. The issue of onset of action is not unique to antidepressants. Although the antihypertensive action of beta blockers has an instantaneous onset, in cases of severe hypertension an additional antihypertensive effect develops in the course of treatment (page 369 in: Cruickshank and Prichard 1994). Could the onset of the clinical action of antidepressants also be related to the severity of the condition to be treated?

References:

Bennabi D, Vandell P, Papaxanthis C, Pozzo T, Haffen E. Psychomotor retardation in depression: a systematic review of diagnostic, pathophysiologic, and therapeutic implications. *Biomed Res Int.* 2013;2013:158746.

Cruickshank JM, Prichard BNC. *Beta-blockers in Clinical Practice.* Edinburgh: Churchill Livingstone; 1994.

Greden JF, Albala AA, Smokler IA, Gardner R, Carroll BJ. Speech pause time: a marker of psychomotor retardation among endogenous depressives. *Biol Psychiatry*, 1981;16:851-9.

Greden JF, Carroll BJ. Decrease in speech pause times with treatment of endogenous depression. *Biol Psychiatry*, 1987;15:575-87.

Hoffmann GM, Gonze JC, Mendlewicz J. Speech pause time as a method for the evaluation of psychomotor retardation in depressive illness. *Br J Psychiatry*, 1985;146:535-8.

Szabadi E, Bradshaw CM. Speech pause time: behavioural correlate of mood. *Am J Psychiatry*, 1983;140:265.

Szabadi E, Bradshaw CM, Besson JAO. Elongation of pause-time in speech: a simple, objective measure of motor retardation in depression. *Br J Psychiatry*, 1976;129:592-7.

August 13, 2015

Martin M. Katz's reply to Elemer Szabadi's comment

Dr. Szabadi is responding to the question of whether the findings from controlled clinical studies that show the large majority of treatment-responsive patients to significantly improve within the first two weeks of antidepressant treatment applies to individual patients. Don Klein raised that question in his earlier critique. Dr. Szabadi summarizes results from studies conducted by several investigators, using an innovative method of determining onset, i.e., associated speech pause time, to show that in many cases, patients who improve on this measure early, go on to recover with treatment. He makes a good case for early onset, but one would have to examine these studies to ensure that his studies cover a significant number of cases.

Nevertheless, these results are certainly in accord with Kuhn's initial report identifying imipramine as an "antidepressant," in which he, as other well-known clinicians, have been quick to report that when a patient responds to six weeks of treatment, early improvement within the first week of treatment, was a frequent occurrence.

More telling from the scientific vantage is that finding that "early response" is a highly frequent occurrence, significantly greater than occurs with placebo or non-effective drugs, is not simply a statistic. It is a finding that could only occur if significantly more patients demonstrated this response than those who did not. The studies by Stassen et al. (1996) and Szegedi et al. (2009) report significant results on this issue based on samples which number in the thousands.

On the issue of the possible association with severity of the disorder that Dr. Szabadi raises, my own experience in our Collaborative program (Maas et al. 1980, 1984) with this issue provides very relevant results. In that study, 100+ severely depressed patients, all severe enough to be hospitalized, were assembled. Each patient was individually examined by an experienced clinician, rated weekly, then categorized on the basis of all behavioral measures, as to whether he/she was "fully recovered" through "no change." The final results of treatment showed 65% of patients "markedly improved."

This finding on "individual patients" was to be expected. What was surprising was that by the end of two weeks of treatment, 50% of patients were categorically judged by experienced clinicians, as "recovered." This helps to explain our reported results on early onset (Katz et al. 1987) and clearly demonstrates that it was based not simply on average response during the first two weeks, but on marked change toward recovery at this early point, by no less than 50% of the sample of patients.

References:

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

Kielholz P, Poldinger W. Die behandlung endogener depressionen mit psychopharmaka. *Dt med Wschr.* 1968;93:701-4.

Maas JW, Koslow S, Davis J, Katz MM, Mendels J, Robins E, Stokes P, Bowden C. Biological component of the NIMH-Clinical Research Branch Collaborative Program on the Psychobiology Depression. *Psychological Medicine*, 1989;10:759-76.

Maas JW, Koslow S, Katz MM, Gibbons RL, Bowden CL, Robins E, Davis JM. Pretreatment neurotransmitter metabolites and tricyclic antidepressant drug response. *Am J Psychiatry*, 1984;141:1159-71.

Stassen HH, Angst J, Delini-Stula A. Delayed onset of action of antidepressant drugs? Survey of recent results. *Eur Psychiatry*, 1996;12:166-76.

Szegedi A, Jansen WT, van Wugenburg AP. Early improvement in the first two weeks as predictors of treatment outcome in patients with major depressive disorder: a meta-analysis including 6,562 patients. *J clin Psychiatry*, 2009;70:344-53.

October 29, 2015

Donald Klein's reply to Elemer Szabadi's comment

The debate between Martin Katz and me centers on the question of how soon after antidepressant treatment is initiated does specific treatment benefit become manifest. The relevant measures used were weekly psychiatric scales, in placebo controlled clinical trials.

Dr. Szabadi (Szabadi et al. 1976) proposes an alternative measure: decreases in the speech pause time during automatic counting. His pilot study of four subjects during antidepressant treatment was promising. The article does not refer to the "small group" that showed effects during the first treatment week. Since there were only four subjects, one wonders at the size of this subgroup.

However, several studies found similar findings relating decreased speech pause time to improvement in depression. One mentioned that agitated depressions showed opposite effects. I

was unable to find any placebo controlled clinical trial that used speech pause time as a dependent measure and I would appreciate such references.

Dr. Szabadi's simple, objective observation has been neglected. Any contribution to the Katz-Klein disagreement depends on its relative value to the usual scales for detecting clinical improvement. Conceivably, it might be measuring the degree of retardation, which is clinically difficult to evaluate but is an important component of remission.

If it were shown that only those drug-treated patients with marked decreases in pause time during the first treatment week went on to substantial clinical improvement, this would be in harmony with Marty's views. Contrariwise, if many patients showed a later onset of substantial maintained improvement that would argue against the stand that early improvement can serve as a surrogate for eventual improvement.

I think the data generated by scales have already invalidated Katz's hypotheses. Nevertheless, such promising alternative measures are, in principle, welcome. Funding their evaluation is another issue.

October 22, 2015

Martin M. Katz's reply to Elemer Szabadi's comment

Dr. Szabadi is responding to the question of whether the findings from controlled clinical studies that show the large majority of treatment-responsive patients to significantly improve within the first two weeks of antidepressant treatment applies to individual patients. Don Klein raised that question in his earlier critique. Dr. Szabadi summarizes results from studies conducted by several investigators, using an innovative method of determining onset, i.e., associated speech pause time, to show that in many cases, patients who improve on this measure early, go on to recover with treatment. He makes a good case for early onset but one would have to examine these studies to ensure that his studies cover a significant number of cases.

Nevertheless, these results are certainly in accord with Kuhn's initial report identifying imipramine as an "antidepressant", in which he, as other well-known clinicians, have been quick to report that when a patient responds to six weeks of treatment, early improvement within the first week of treatment, was a frequent occurrence.

More telling from the scientific vantage is that finding that “early response” is a highly frequent occurrence, significantly greater than occurs with placebo or non-effective drugs, is not simply a statistic. It is a finding that could only occur if significantly more patients demonstrated this response than those who did not. The studies by Stassen et al (1996) and Szegedi et al (2009) report significant results on this issue based on samples which number in the thousands.

On the issue of the possible association with severity of the disorder that Dr. Szabadi raises, my own experience in our Collaborative program (Maas et al 1980.1984) with this issue provides very relevant results. In that study, 100+ severely depressed patients, all severe enough to be hospitalized, were assembled. Each patient was individually examined by an experienced clinician, rated weekly, then categorized on the basis of all behavioral measures, as to whether he/she was “fully recovered” through “no change”. The final results of treatment showed 65% of patients “markedly improved”.

This finding on “individual patients” was to be expected. What was surprising was that by the end of two weeks of treatment, 50% of patients were categorically judged by experienced clinicians, as “recovered”. This helps to explain our reported results on early onset (Katz et al 1987) and clearly demonstrates that it was based not simply on average response during the first two weeks, but on marked change toward recovery at this early point, by no less than 50% of the sample of patients.

References:

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

Kielholz P, Poldinger W. Die behandlung endogener depressionen mit psychopharmaka. *Dt med Wschr* 1968; 93:701-4.

Maas JW, Koslow S, Davis J, KatzMM, Mendels J, Robins E, Stokes P, Bowden C. Biological component of the NIMH-Clinical Research Branch Collaborative Program on the Psychobiology Depression. *Psychological Medicine* 1989; 10:759-76.

Maas JW, Koslow S, Katz MM, Gibbons RL, Bowden CL, Robins E, Davis JM. Pretreatment neurotransmitter metabolites and tricyclic antidepressant drug response. *Am J Psychiatry* 1984; 141:1159-71.

Stassen HH, Angst J, Delini-Stula A. Delayed onset of action of antidepressant drugs? Survey of recent results. *Eur Psychiatry* 1996; 12: 166-76.

Szegedi A, Jansen WT, van Wugenburg AP. Early improvement in the first two weeks as predictors of treatment outcome in patients with major depressive disorder: a meta-analysis including 6,562 patients. *J clin Psychiatry* 2009; 70:344-53.

October 29, 2015

June 24, 2019

January 16, 2020