ORIGINAL RESEARCH

Key Words: onset of action, antidepressant, onset efficacy, study design, study methodology, noninferiority

Addressing Methodological Issues in Studying Antidepressant Onset Efficacy Using Prespecified Sensitivity Analyses

By Ilya Lipkovich, PhD, Craig H. Mallinckrodt, PhD, Apurva Prakash, BA, and Andrew A. Nierenberg, MD

ABSTRACT *~* ***Objectives:*** *Assessing antidepressant onset efficacy presents substantial methodological challenges. Most assessments of onset efficacy are based on post hoc analyses. This article presents a case study of a prospectively designed trial to compare antidepressant onset efficacy.* ***Experimental Design:*** *The current study design was compared with previously published criteria for an ideal onset of action study. Prospectively defined sensitivity analyses were used to assess whether results of the present study were influenced by the assumptions and decisions necessary to implement this study.* ***Principal Observations:*** *The study achieved its primary objective of establishing noninferiority between SNRI and SSRI. Sensitivity analyses suggested that results from the primary analysis were not influenced by patient population, outcome measure, cut-off for defining clinically meaningful sustained improvement, or analytical method. Although not all limitations could be addressed, appropriate conclusions could be drawn. For example, the study tested only one fixed dose of each drug; hence, conclusions are limited to those dosages and cannot be extrapolated across the entire dose ranges, as would be possible in the ideal study.* ***Conclusion:*** *This article illustrates that prospectively designed studies (as opposed to retrospective comparisons) can be implemented and sensitivity analyses can be used to address concerns regarding assumptions and arbitrary decisions.* Psychopharmacology Bulletin. *2008;41(2):40–54.*

INTRODUCTION

Antidepressant medications typically exhibit a delay of at least 2 weeks following the start of therapy before patients experience clinically relevant improvement.[1–3] During this latency period, patients remain symptomatic and functionally impaired, with an associated risk for suicide and morbidity, a prolonged reduction in quality of life, and a continued loss of work productivity. In contrast, antidepressants with a more rapid onset of action may offer potential

Dr. Lipkovich, Dr. Mallinckrodt, and Ms. Prakash are affiliated with Lilly Research Laboratories, Eli Lilly and Company, Indianapolis, IN. Dr. Nierenberg is affiliated with Depression Clinical and Research Program, Massachusetts General Hospital, Boston, MA.

To whom correspondence should be addressed: Craig H. Mallinckrodt, PhD, Eli Lilly and Company, Lilly Corporate Center DC 0734, Indianapolis, IN 46285; E-mail: cmallinc@lilly.com

benefits. Therefore, the development of new pharmacotherapies, with rapid onset of action, is an active area of research.[2,4–11]

Assessments of onset of action have generally been based on retrospective analyses.[12–14] Although post hoc analyses of existing data may provide an initial indication of differential times to onset, a prospectively designed study with an a priori-specified definition of onset of action is generally preferred.[15] Unfortunately, few adequately powered, prospectively designed onset of action studies have been undertaken.[16–18] This may be the result of the methodological challenges inherent in designing such trials, which has also led to an extensive body of literature in this area.[12,15,19–21]

Leon et al.[15] addressed these methodological challenges in proposing a design for an "ideal" trial to study differential onset of action between two active drugs. The proposed design called for the enrollment of 1,750 patients, leading the authors to admit that "it would be impractical, if not impossible, to conduct a study based on our design." Rather than abandoning the opportunity to prospectively study onset, our approach was to design the best trial possible and then assess via sensitivity analyses how the assumptions and decisions necessary for implementing the trial may have influenced results.

The purpose of this investigation was to outline how our trial differed from the ideal trial and to present results of the sensitivity analyses that assessed how these differences and other assumptions influenced results.

MATERIALS AND METHODS

Study Design

The study design along with primary efficacy and safety results have been described elsewhere.[22] Key details necessary to understand onset of action assessments are noted in this section. This was a randomized, double-blind, placebo- and active comparator-controlled study conducted at 36 sites (psychiatric clinical settings) in the United States. Patients meeting entry criteria were randomly assigned (2:2:1 ratio) to 1 of 2 active antidepressants (referred to here as SNRI and SSRI), or placebo, respectively. After acute treatment, patients continued on double-blind therapy for an additional 6 months. Other relevant details, and how the design of this study compared to the ideal study as defined by Leon et al., are summarized in Table 1.

The study incorporated a double-blind, variable expected-duration placebo lead-in period to blind patients and investigators to the start of active therapy. Clinic visits were scheduled weekly during the blinded lead-in and at weeks 1, 2, 3, 4, 6, and 8 while patients were on active therapy.

TABLE 1

COMPARISON OF "IDEAL" STUDY CRITERIA WITH THOSE EMPLOYED IN THE PRESENT STUDY

	"IDEAL" STUDY[15]	PRESENT STUDY
Study design	Randomized, double-blind, placebo-controlled	Randomized, double-blind, placebo-controlled
Duration	8 weeks	8 weeks
Study drugs	2 active drugs	2 active drugs
Dosing regimens	2 per drug	1 per drug
Definition of onset	Prospective	Prospective
Onset criteria	30% reduction in $HAMD_{17}$ Total score	20% reduction in $HAMD_{17}$ Maier subscale score (primary); 30% reduction in $HAMD_{17}$ Maier subscale score (secondary)
Timing of onset	By week 4	Assessed at week 2
Duration of improvement	Sustained to study endpoint	Sustained to study endpoint
Frequency of assessment	Twice per week	Weekly (primary); twice per week (secondary)
Powering	≥80%	80%
Patients	Meet DSM-IV criteria for MDD	Meet DSM-IV criteria for MDD
Baseline illness severity	$HAMD_{17}$ Total score ≥20	MADRS Total score ≥22, and CGI-S score ≥4
Age	25–50 years	≥18 years
Previous treatment	Naïve to treatment with study drugs	Patients excluded for previous lack of response to SNRI or SSRI
Patient/investigator blinding	(i) Variable-duration placebo lead-in; (ii) investigators blinded to study visit	(i) Variable-duration placebo lead-in; (ii) Investigators not blinded to study visit
Data analysis	Mixed-effects model	Mixed-effects model

Abbreviations: $HAMD_{17}$, 17-item Hamilton Depression Rating scale; DSM-IV, Diagnostic and Statistical Manual of Mental Disorders; MDD, major depressive disorder; MADRS, Montgomery-Asberg Depression Rating Scale; CGI-S, clinical global impressions-severity of illness.

Lipkovich, Mallinckrodt, Prakash, et al. *Psychopharmacology Bulletin*. Vol. 41. No. 2. 2008.

Study Patients. Study participants were adult outpatients at least 18 years of age. All patients met diagnostic criteria for major depressive disorder (MDD) as defined in the *Diagnostic and Statistical Manual of Mental Disorders* (DSM-IV).[23] The diagnosis of MDD was confirmed by the Mini International Neuropsychiatric Interview (MINI).[24] Patients were required to have a Montgomery-Asberg Depression Rating Scale (MADRS)[25] Total score ≥22 and a Clinical Global Impressions-Severity

of Illness (CGI-S)[26] score ≥4 at the two screening study visits. Other exclusion criteria were similar to those used in other studies of SNRI. All study patients provided written informed consent prior to study entry and in accordance with the principles of the Declaration of Helsinki.

Primary Objective. The primary objective of the study was to compare the onset of antidepressant efficacy for the two drugs by testing the hypothesis that the percentage of patients treated with SNRI who had clinically meaningful and sustained improvements in core depressive symptoms was not inferior to (at least as good as) SSRI at week 2. Investigators and patients were blinded to the timing of the primary assessment.

Primary Efficacy Measure. Onset of action was defined as achieving a 20% decrease from baseline on the 17-item Hamilton Depression Rating scale ($HAMD_{17}$) Maier subscale,[27] and maintaining or exceeding that reduction at all subsequent visits throughout acute treatment. Efficacy measures were assessed at the regularly scheduled clinic visits. In addition, the Hospital Anxiety and Depression Scale (HADS)[28] was obtained via an Interactive Voice Response (IVR) telephone system twice-weekly until week 4 of active therapy.

The Maier subscale includes $HAMD_{17}$ Items 1 (depressed mood), 2 (guilt), 7 (work and activities), 8 (retardation), 9 (agitation), and 10 (psychic anxiety). Faries et al.[29] reported that subscales of the $HAMD_{17}$ might provide more precise discrimination of antidepressant effects than the Total score. The use of subscales may be especially appropriate for assessing onset of efficacy, as it has been hypothesized that some early responses to antidepressant medication may be attributable to improvements in sleep rather than core depressive symptoms.[30]

For this study, the 20% sustained improvement on the Maier subscale definition of onset was supported by the general onset methodology literature[15,20] and specifically by previous results from SNRI.[31]

Study Powering and Analytical Methods. The powering of the study was based on an anticipated enrollment of 675 patients. This sample size was estimated to have 80% power for the primary analysis that tested the noninferiority of SNRI against SSRI. Assumptions underlying the power estimate included the following: (i) the use of a one-sided 97.5% (two-sided 95%) confidence interval (CI); (ii) 40% of SSRI-treated patients would meet onset criteria at the primary endpoint of week 2; (iii) the maximum disadvantage of SNRI compared with SSRI that was not considered meaningful (the delta margin, defined by the

lower bound of the CI) was −10%; and (iv) the true advantage of SNRI at week 2 would be ~2.5%.

Onset was defined as a binary outcome (yes/no) with noninferiority based on the intent-to-treat population using a one-sided 97.5% (two-sided 95%) CI. The primary efficacy analysis of onset used a pseudo-likelihood-based categorical repeated measures approach (Categorical MMRM), similar to that discussed by Leon.[32] This analysis included the fixed, categorical effects of treatment, investigator, visit, and treatment-by-visit interaction, as well as the continuous, fixed covariate of baseline score. The analysis also included a logit link function and assumed binomial errors. An unstructured covariance matrix was found to provide the best fit to the data (based on an a priori-specified algorithm using Akaike's information criterion) and was used to model the within-patient errors. The Kenward-Roger approximation was used to estimate denominator degrees of freedom. The analysis was implemented using the GLIMMIX macro in SAS.

Secondary comparisons of the primary analysis were defined a priori to include contrasts for the main effect of treatment and the treatment-by-visit interaction, with tests considered statistically significant if the two-sided *P*-value was ≤.05.

Sensitivity Analyses

A series of secondary analyses were prespecified to assess the robustness of the primary result against the assumptions and arbitrary decisions inherent to the primary analysis. These analyses included several domains:

1. Patient population,
2. Cut-off for defining clinically meaningful sustained improvement,
3. Efficacy measure,
4. Assessment frequency, and
5. Analytic method

Patient Population. In addition to the all-patient (intent-to-treat) cohort, the primary onset analysis was conducted on three additional patient populations:

i. Qualified: Patients with high placebo response during the double-blind placebo lead-in (≥30% reduction in $HAMD_{17}$ Total score) were excluded.
ii. Per-protocol: Patients with significant protocol violations were excluded. Significant protocol violations focused primarily on the use of excluded medications that would likely influence efficacy results.

iii. Qualified and Per-protocol: Patients with high placebo response or significant protocol violations were excluded.

Cut-Off. Although the primary definition of onset utilized a 20% sustained improvement, analyses were repeated using 30% sustained improvements as well.

Efficacy Measure. The primary analysis utilized the $HAMD_{17}$ Maier subscale. Secondary analyses used the Hamilton Anxiety scale (HAMA) Psychic subscale and the HADS Depression subscale.

Assessment Frequency. Previous investigations of onset methodology have recommended twice-weekly visits.[15] However, a more recent study concluded that a traditional (weekly) assessment schedule was reasonable when comparing treatment group differences in onset of action using a categorical repeated measures approach.[33] In this study, the primary analysis was applied to Clinical Report Form (CRF) data obtained at weeks 1, 2, 3, 4, 6, and 8. However, similar analyses were also applied to patient-rated (IVR) data from the HADS that was obtained twice-weekly through week 4.

Analytic Method. Although previous research has provided justification for the use of the Categorical MMRM analysis as the primary approach, it has been proposed that no single analysis can be fully justified on a priori grounds when an appreciable percentage of patients have missing data. Lipkovich et al.[34] suggested that when the probability of dropouts depends on previously observed outcomes, the Categorical MMRM analysis may lead to biased inference since, unlike similar approaches in continuous data, a pseudo-likelihood approach is used rather than a direct-likelihood approach. Therefore, alternative approaches may provide better estimates of treatment difference in the categorical outcome. Several alternative analytic approaches were implemented:

i. Kaplan-Meier survival analysis of time to onset.
ii. A logistic regression analysis of probability of onset at week 2 with missing observations imputed via last observation carried forward (LOCF).
iii. Repeated measures analyses of probability of onset using a generalized estimating equations (GEE) approach.[35]
iv. A weighted GEE approach with weights determined at every time point by the inverse of estimated probability of remaining on treatment.[36]

v. A multiple imputation (MI) analysis of the repeated measures data with restrictive and inclusive strategies[37] for constructing imputation models for the underlying continuous score ($HAMD_{17}$ Maier subscale) and categorical repeated measures (GEE) for analysis of probability of onset in completed datasets.

RESULTS

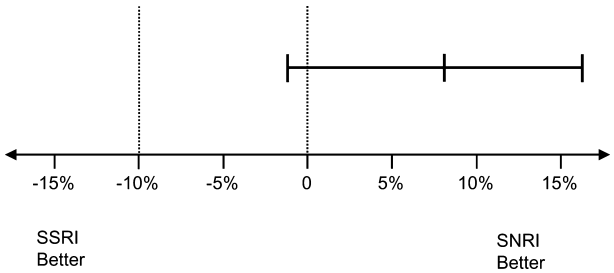
Results from the primary outcome measure suggested that SNRI was not inferior to SSRI in the onset of antidepressant action. Probabilities of meeting onset criteria (20% sustained reduction in $HAMD_{17}$ Maier subscale score at week 2) were 42.6% vs. 35.2% for SNRI and SSRI, respectively (treatment difference = 7.4%; 95% CI, −1.3% to 16.2%; $P = .097$; Figure 1). The probability of meeting onset criteria among placebo-treated patients (21.5%) was significantly lower than that for either active drug group ($P < .001$ for SNRI vs. placebo; $P = .008$ for SSRI vs. placebo).

In the sensitivity analyses, the significance of the noninferiority outcome was unaffected by choice of patient group (all; qualified; per-protocol; or qualified and per-protocol; Table 2). Noninferiority of SNRI was also demonstrated using a more stringent cut-off point (30% sustained reduction in $HAMD_{17}$ Maier subscale score) and other efficacy outcomes (HADS Depression subscale; Table 2).

The percentage of SNRI-treated patients meeting sustained onset criteria was significantly higher than that in the placebo group at each

FIGURE 1

RESULT OF THE PRIMARY OUTCOME: NONINFERIORITY TEST FOR SNRI VERSUS SSRI. PROBABILITIES OF MEETING ONSET CRITERIA (20% SUSTAINED REDUCTION IN $HAMD_{17}$ MAIER SUBSCALE SCORE AT WEEK 2) FOR SNRI- AND SSRI-TREATED PATIENTS WERE 42.6% VS. 35.2%, RESPECTIVELY (TREATMENT DIFFERENCE = 7.4%; 95% CONFIDENCE INTERVAL, −1.3% TO 16.2%; $P = .097$). (ABBREVIATION: $HAMD_{17}$, 17-ITEM HAMILTON DEPRESSION RATING SCALE).



Lipkovich, Mallinckrodt, Prakash, et al. *Psychopharmacology Bulletin*. Vol. 41. No. 2. 2008.

TABLE 2

SUMMARY OF ONSET OF ACTION OUTCOMES

ONSET CRITERION	PATIENTS	PROBABILITY OF ONSET AT WEEK 2 (%) SNRI ($N = 262$)	SSRI ($N = 267$)	PLACEBO ($N = 135$)	NONINFERIORITY MARGIN, (%) (CI)
20% sustained reduction in $HAMD_{17}$ Maier subscale score	All	42.6*	35.2*	21.5	7.4 (−1.3%, 16.2%)
20% sustained reduction in $HAMD_{17}$ Maier subscale score	Qualified	44.0*	37.5*	22.0	6.5 (−3.3%, 16.4%)
20% sustained reduction in $HAMD_{17}$ Maier subscale score	Per-Protocol	45.1*	36.5*	21.6	8.6 (−0.5%, 17.7%)
20% sustained reduction in $HAMD_{17}$ Maier subscale score	Qualified and Per-Protocol	46.2*	38.5*	22.0	7.7 (−2.4%, 17.9%)
30% sustained reduction in $HAMD_{17}$ Maier subscale score	All	29.1*	22.9*	10.8	6.2 (−1.6%, 14.1%)
20% sustained reduction in HADS Depression subscale score	All	13.4	11.6	8.9	1.8 (−3.2%, 6.8%)

Abbreviations: $HAMD_{17}$, 17-Item Hamilton Depression Rating scale; HADS, Hospital Anxiety and Depression Scale.
*$P < .01$ vs. placebo.

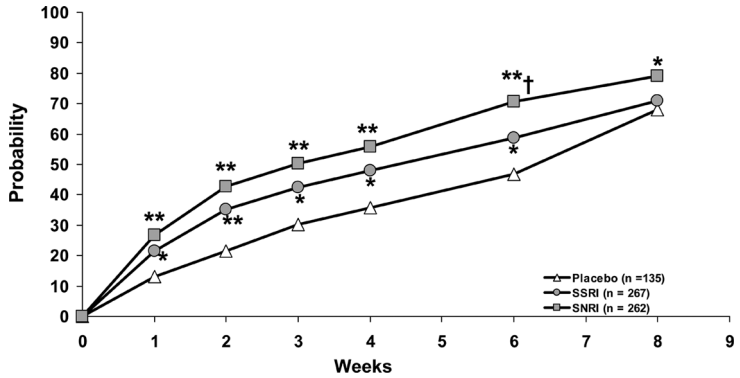
Lipkovich, Mallinckrodt, Prakash, et al. *Psychopharmacology Bulletin*. Vol. 41. No. 2. 2008.

visit during the acute-treatment phase (Figure 2). The probability of achieving onset was significantly higher in the SSRI group, when compared with the placebo group, at all visits except endpoint (week 8; Figure 2).

Using the main effect of treatment (pooling all visits), a significantly greater proportion of SNRI-treated patients met onset criteria (20% sustained reduction in $HAMD_{17}$ Maier subscale score) when compared with those receiving SSRI ($P = .026$; Table 3). Both SNRI ($P < .001$) and SSRI ($P = .018$) groups were significantly superior to placebo on this measure. Results from the main effect of treatment were unaffected by the choice of patient group (Table 3). Both drugs were significantly superior to placebo on the main effect of treatment when the specified

FIGURE 2

VISITWISE PROBABILITIES OF ACHIEVING 20% SUSTAINED REDUCTION IN $HAMD_{17}$ MAIER SUBSCALE SCORE FOR PATIENTS RECEIVING SNRI, SSRI, OR PLACEBO. *$P \leq .05$ VS. PLACEBO; **$P \leq .01$ VS. PLACEBO; †$P \leq .05$ VS. SSRI. (ABBREVIATION: $HAMD_{17}$, 17-ITEM HAMILTON DEPRESSION RATING SCALE).



Lipkovich, Mallinckrodt, Prakash, et al. *Psychopharmacology Bulletin*. Vol. 41. No. 2. 2008.

onset criteria were a 30% sustained reduction in $HAMD_{17}$ Maier subscale score.

Neither of the active drugs was superior to placebo on the main effect of treatment when the specified onset criterion was a 20% sustained reduction in HADS Depression subscale score (Table 3). However, in mean change analyses (not shown), both drugs produced significant improvement when compared with placebo at the first assessment (following 3 days of treatment) and at most subsequent assessments.

Using Kaplan-Meier analysis, the time to onset (20% reduction in $HAMD_{17}$ Maier subscale score) for SNRI-treated patients (median = 23 days) was significantly shorter than that for patients receiving SSRI (median = 41 days; $P = .032$ from log-rank test) or placebo (median = 55 days; $P < .001$ from log-rank test; Table 4). The difference between SSRI and placebo in time to onset was not statistically significant ($P = .087$ from log-rank test). Results from additional Kaplan-Meier analyses of time to onset using different onset criteria are presented in Table 4.

Logistic regression results based on LOCF for weeks 2–8 were consistent with results from the primary a priori analysis (Categorical MMRM). Based on LOCF analyses, the advantage for SNRI (treatment difference in probability of onset at week 2) was 6.6% (95% CI: −1.3, 14.4); $P = .10$. Categorical MMRM yielded an estimate of advantage for SNRI of 7.4% (95% CI: −1.3, 16.2); $P = .10$. Results using other analytic methods (GEE, weighted GEE, and MI using restrictive and inclusive strategies) were consistent with the categorical MMRM results showing estimated treatment advantages for SNRI ranging from

TABLE 3

SUMMARY OF ONSET OF ACTION OUTCOMES (MAIN EFFECT OF TREATMENT)

		P-VALUE FOR MAIN EFFECT OF TREATMENT		
ONSET CRITERION	PATIENTS	SNRI VS. PLACEBO	SSRI VS. PLACEBO	SNRI VS. SSRI
20% sustained reduction in $HAMD_{17}$ Maier subscale score	All	<.001	.018	.026
20% sustained reduction in $HAMD_{17}$ Maier subscale score	Qualified	<.001	.022	.046
20% sustained reduction in $HAMD_{17}$ Maier subscale score	Qualified and Per-Protocol	<.001	.025	.015
30% sustained reduction in $HAMD_{17}$ Maier subscale score	All	<.001	.007	.063
20% sustained reduction in HADS Depression subscale score	All	.052	.120	.644

Abbreviations: $HAMD_{17}$, 17-Item Hamilton Depression Rating scale; HADS, Hospital Anxiety and Depression Scale.

Lipkovich, Mallinckrodt, Prakash, et al. *Psychopharmacology Bulletin*. Vol. 41. No. 2. 2008.

TABLE 4

SUMMARY OF MEDIAN TIME TO ONSET

	MEDIAN TIME TO ONSET (DAYS)		
ONSET CRITERION	SNRI	SSRI	PLACEBO
20% sustained reduction in $HAMD_{17}$ Maier subscale score	23*,**	41	55
30% sustained reduction in $HAMD_{17}$ Maier subscale score	42*	47	56
20% sustained reduction in HADS Depression subscale score	32	35	35

Abbreviations: $HAMD_{17}$, 17-Item Hamilton Depression Rating scale; HADS, Hospital Anxiety and Depression Scale.
*$P < .01$ vs. placebo (log-rank test). **$P <. 05$ vs. SSRI (log-rank test).

Lipkovich, Mallinckrodt, Prakash, et al. *Psychopharmacology Bulletin*. Vol. 41. No. 2. 2008.

7.5 to 8.1% with the lower bound of the 95% CI ranging from −1.5 to −0.5 and associated *P*-values ranging from .10 to .07.

DISCUSSION

The results of this investigation suggest that the comparison of SNRI with SSRI in onset of effect was robust to key decisions and assumptions. Specifically, the primary outcome, demonstrating noninferiority of SNRI versus SSRI, was unaffected by choice of patient population,

cut-off threshold, efficacy measure, or analytic technique. The impact of assessment frequency was less clear due to the confounding of assessment measure with assessment frequency.

Some specific aspects of how this study differed from the ideal study warrant further discussion. Leon et al. proposed that onset of symptom improvement should be monitored using the $HAMD_{17}$ Total score. However, antidepressants with sedating properties may have a rapid effect on the $HAMD_{17}$ subscale items related to insomnia (items 4, 5, and 6), but might not improve the core mood symptoms of depression.[30] Thase[21] noted that "although advantages in improving selected symptom constellations more rapidly might be of benefit for some patients, the objective is to identify drugs with a more rapid action on the core symptoms of depression." Thus, using the Maier subscale,[27] which focuses on the core emotional symptoms of depression, as the primary outcome in this investigation was reasonable and potentially preferable in using the $HAMD_{17}$ Total score.

Leon et al. proposed the use of a 30% reduction from baseline in symptom severity as a measure of onset of action. Although this definition has been employed by other researchers,[38] the use of a 20% reduction from baseline also has literature precedent.[39–42] In this study, the 20% threshold was employed as the primary outcome, but conclusions were robust to the use of either cut-off.

Universal agreement exists regarding the importance of basing onset assessments on sustained treatment responses because early improvements are of little value unless they are sustained.[15,30,43–45] Therefore, this study required improvements to be maintained throughout the acute-treatment period.

Leon et al.[15] recommended the use of twice-weekly assessments during the initial stages of an onset of action study. This study utilized weekly patient visits. Frequent visits and assessments may influence patient recruitment, retention, the patient population, and other logistics, as well as cost and placebo response. Such a trial has never been done, at least not on the scale needed to accurately compare active drugs on widely accepted and meaningful endpoints.

A recent study suggested that a standard assessment schedule was sufficient to compare treatment groups in onset of effect when a mixed-effects model repeated measures approach is used; however, when survival analyses are employed, or when the focus is on time to onset for individual patients, a more frequent assessment schedule may be beneficial.[33] This study included twice-weekly assessments using the patient-rated HADS via an IVR system as a secondary outcome in addition to the clinician-rated $HAMD_{17}$ assessments obtained according to a traditional assessment schedule (weeks 1, 2, 3, 4, 6, and 8).

Results from the traditional assessment schedule clearly differentiated both active drugs from placebo on a number of measures, thereby establishing the assay sensitivity of the study. Several secondary analyses that examined the entire acute phase rather than a single time point yielded statistically significant advantages for SNRI over SSRI, further substantiating the primary results from the traditional assessment schedule. Results from the twice-weekly assessments were directionally consistent with those obtained from the standard schedule, but were less robust and did not achieve statistical significance.

In assessing differences in time to onset between active drugs, it has been emphasized that between-group differences should not only be statistically significant but also clinically relevant. Leon et al. noted that time to onset differences of a few days would be of little value, whereas a difference of a week or more may have a significant clinical impact.[15] In this study, the advantage of SNRI over SSRI in median time to onset was 18 days (23 days vs. 41 days, respectively). However, based on the Categorical MMRM analysis, SSRI attained the same percentage of patients meeting onset as SNRI at the next assessment time. It is likely that the Kaplan-Meier analysis over-estimated the difference in median times as a consequence of not frequently assessing patients. SSRI almost achieved onset at week 4, but as there was no another assessment until week 6, median onset could not be achieved until that time. The ability of the Categorical MMRM analysis to yield useful estimates of onset independent of the assessment schedule has been noted and is a factor in making onset assessments feasible from traditional assessment schedules.[33] Therefore, in this study, it would be more accurate to base judgments regarding the clinical relevance of differences on results from the Categorical MMRM approach.

It is perhaps not surprising that results were insensitive with respect to the different methods of handling missing data, given that the primary analysis was based on evaluations at week 2, and very few patients had dropped out by that early time point.

Despite the extensive sensitivity analyses, these results should be considered in light of remaining limitations. Specifically, the magnitude of the chosen noninferiority margin may be debated; a difference of 5% may be preferable to the stated margin of 10%. However, given that SNRI substantially surpassed the prespecified margin of 10% and the alternative 5% margin mitigates concerns surrounding this issue.

Although the use of only one dosing regimen of each active drug dramatically reduced the required sample size of the study, a tradeoff exists in terms of the generalizability of the study results. This study compared onset of SNRI at a fixed dose versus SSRI at a fixed dose. Therefore, conclusions drawn from these data are specific to the doses tested and

cannot be extrapolated to SNRI and SSRI across their approved dose ranges.

This study utilized inclusion and exclusion criteria similar to those employed in the registration studies and other clinical trials of SNRI rather than the more stringent criteria recommended by Leon et al.[15] This decision reflects almost the ever-present tradeoffs between internal validity and external validity. The more restrictive patient population could have reduced heterogeneity and provided a clearer signal of the differences between treatments. However, the inclusion of these additional groups (e.g., elderly patients, those with previous depressive episodes) may also make these results more applicable to clinical practice, where elderly and recurrent patients are common.

Leon et al.[15] proposed that investigators and assessors should be blinded to the study visit in an onset of action clinical trial. This feature was not incorporated into the design of this study. However, blinding investigators and patients to the timing of initiation of active study drug may have helped reduce the expectation bias.

CONCLUSIONS

Studying onset of antidepressant action poses formidable methodological challenges. This article illustrates, however, that prospectively designed studies (as opposed to retrospective comparisons) can be implemented and that sensitivity analyses can be used to address concerns regarding assumptions and arbitrary decisions.

Based on these results, we offer the following recommendations when developing future studies to assess antidepressant onset efficacy: (1) A unidimensional subscale assessing core depressive symptoms (such as the $HAMD_{17}$ Maier subscale) may be able to better distinguish differing drug effects when compared with the multidimensional $HAMD_{17}$ Total score. (2) A traditional assessment schedule (weekly visits) is adequate for assessing treatment group differences when used in conjunction with a categorical repeated measures analytic approach and does not increase the burden of implementing the study, the study cost, or the risk of increased placebo response (as would be more likely for a study using more frequent study visits). (3) Onset of action can be based on 20–30% initial improvement, and the additional requirement that the improvement be sustained or exceeded throughout acute treatment may better differentiate drug effects from placebo. ✤

ACKNOWLEDGMENTS

The authors thank Andrew C. Leon, Ph.D., for reviewing drafts of this manuscript and for providing critical review and feedback. This study was sponsored by Eli Lilly and Company. Dr. Nierenberg is an

employee of Massachusetts General Hospital, is a consultant for Eli Lilly and Company, Shire, GSK, Innapharma, Genaissance, Sepracor, Novartis, and Somerset, and has received grant support from Eli Lilly and Company, Wyeth, GSK, BMS, Cyberonics, Lichtwer, Pfizer, Cederroth, Forest, Janssen, NARSAD, and NIMH, and has also received honoraria from Eli Lilly and Company, Wyeth, GSK, and Cyberonics. Drs. Lipkovich and Mallinckrodt and Ms. Prakash are stock shareholders and full-time employees of Eli Lilly and Company.

REFERENCES

1. Nierenberg AA, Farabaugh AH, Alpert JE, Gordon J, Worthington JJ, Rosenbaum JF, et al. Timing of onset of antidepressant response with fluoxetine treatment. *Am J Psychiatry.* 2000;157:1423-1428.
2. Blier P. Pharmacology of rapid-onset antidepressant treatment strategies. *J Clin Psychiatry*. 2001; 62(Suppl 15):12-17.
3. Stahl SM, Nierenberg AA, Gorman JM. Evidence of early onset of antidepressant effect in randomized controlled trials. *J Clin Psychiatry*. 2001;62(Suppl 4):17-23.
4. Adell A, Castro E, Celada P, Bortolozzi A, Pazos A, Artigas F. Strategies for producing faster acting antidepressants. *Drug Discov Today*. 2005;10:578-585.
5. Artigas F, Romero L, de Montigny C, Blier P. Acceleration of the effect of selected antidepressant drugs in major depression by 5-HT1A antagonists. *Trends Neurosci*. 1996;19:378-383.
6. Blier P. Possible neurobiological mechanisms underlying faster onset of antidepressant action. *J Clin Psychiatry*. 2001;62(Suppl 4):7-11.
7. Briner K, Dodel RC. New approaches to rapid onset antidepressants. *Curr Pharm Des*. 1998;4:291-302.
8. Dremencov E, Gispan-Herman I, Rosenstein M, Mendelman A, Overstreet DH, Zohar J, et al. The serotonin-dopamine interaction is critical for fast-onset action of antidepressant treatment: in vivo studies in an animal model of depression. *Prog Neuropsychopharmacol Biol Psychiatry*. 2004;28:141-147.
9. Miller FE. Strategies for the rapid treatment of depression. *Hum Psychopharmacol*. 2001;16:125-132.
10. Montgomery SA. Fast-onset antidepressants. *Int Clin Psychopharmacol*. 1997;12(Suppl 3):S1-S5.
11. Scorza C, Silveira R, Nichols DE, Reyes-Parada M. Effects of 5-HT-releasing agents on the extracellular hippocampal 5-HT of rats. Implications for the development of novel antidepressants with a short onset of action. *Neuropharmacology*. 1999;38:1055-1061.
12. Entsuah R, Derivan A, Kikta D. Early onset of antidepressant action of venlafaxine: pattern analysis in intent-to-treat patients. *Clin Ther*. 1998;20:517-526.
13. Gorman JM, Korotzer A, Su G. Efficacy comparison of escitalopram and citalopram in the treatment of major depressive disorder: pooled analysis of placebo-controlled trials. *CNS Spectr*. 2002;7:40-44.
14. Quitkin FM, Taylor BP, Kremer C. Does mirtazapine have a more rapid onset than SSRIs? *J Clin Psychiatry*. 2001;62:358-361.
15. Leon AC, Blier P, Culpepper L, Gorman JM, Hirschfeld RM, Nierenberg AA, et al. An ideal trial to test differential onset of antidepressant effect. *J Clin Psychiatry*. 2001;62(Suppl 4):34-36.
16. Behnke K, Sogaard J, Martin S, Bäuml J, Ravindran AV, Agren H, et al. Mirtazapine orally disintegrating tablet versus sertraline: a prospective onset of action study. *J Clin Psychopharmacol*. 2003;23: 358-364.
17. Fabre LF. Treatment of depression in outpatients: a controlled comparison of the onset of action of amoxapine and maprotiline. *J Clin Psychiatry*. 1985;46:521-524.
18. Jouvent R, Le Houezec J, Payan C, Mikkelsen H, Fermanian J, Millet V, et al. Dimensional assessment of onset of action of antidepressants: a comparative study of moclobemide vs. clomipramine in depressed patients with blunted affect and psychomotor retardation. *Psychiatry Res*. 1998;79:267-275.
19. Derivan A, Entsuah AR, Kikta D. Venlafaxine: measuring the onset of antidepressant action. *Psychopharmacol Bull*. 1995;31:439-447.
20. Montgomery SA, Bech P, Blier P, Möller HJ, Nierenberg AA, Pinder RM, et al. Selecting methodologies for the evaluation of differences in time to response between antidepressants. *J Clin Psychiatry*. 2002;63:694-699.
21. Thase ME. Methodology to measure onset of action. *J Clin Psychiatry*. 2001;62(Suppl 15):18-21.
22. Nierenberg AA, Greist JH, Mallinckrodt CH, Prakash A, Sambunaris A, Tollefson GD, et al. Duloxetine versus escitalopram and placebo in the treatment of patients with major depressive disorder: onset of antidepressant action, a non-inferiority study. *Curr Med Res Opin*. 2007;23:401-416.

23. American Psychiatric Association. *Diagnostic and Statistical Manual of Mental Disorders, 4th ed*. Washington, DC: American Psychiatric Press; 1994.
24. Sheehan DV, Lecrubier Y, Sheehan KH, Amorim P, Janavs J, Weiller E, et al. The Mini-International Neuropsychiatric Interview (MINI): the development and validation of a structured diagnostic psychiatric interview for DSM-IV and ICD-10. *J Clin Psychiatry*. 1998;59(Suppl 20):22-33.
25. Montgomery SA, Asberg M. A new depression scale designed to be sensitive to change. *Br J Psychiatry*. 1979;134:382-389.
26. Guy W. *ECDEU Assessment Manual for Psychopharmacology, revised 1976*. Rockville, MD: National Institute of Mental Health; 1976.
27. Maier W, Philipp M. Improving the assessment of severity of depressive states: a reduction of the Hamilton Depression Scale. *Pharmacopsychiatry*. 1985;18:114-115.
28. Zigmond AS, Snaith RP. The hospital anxiety and depression scale. *Acta Psychiatr Scand*. 1983;67: 361-370.
29. Faries D, Herrera J, Rayamajhi J, DeBrota D, Demitrack M, Potter WZ. The responsiveness of the Hamilton Depression Rating Scale. *J Psychiatr Res*. 2000;34:3-10.
30. Quitkin FM, Rabkin JD, Markowitz JM, Stewart JW, McGrath PJ, Harrison W. Use of pattern analysis to identify true drug response. A replication. *Arch Gen Psychiatry*. 1987;44:259-264.
31. Quitkin FM, Rabkin JG, Ross D, Stewart JW. Identification of true drug response to antidepressants. Use of pattern analysis. *Arch Gen Psychiatry*. 1984;41:782-786.
32. Leon AC. Measuring onset of antidepressant action in clinical trials: an overview of definitions and methodology. *J Clin Psychiatry*. 2001;62(Suppl 4):12-16.
33. Mallinckrodt CH, Detke MJ, Kaiser CJ, Watkin JG, Molenberghs G, Carroll RJ. Comparing onset of antidepressant action using a repeated measures approach and a traditional assessment schedule. *Stat Med*. 2006;25:2384-2397.
34. Lipkovich I, Duan Y, Ahmed S. Multiple imputation compared with restricted pseudo-likelihood and generalized estimating equations for analysis of binary repeated measures in clinical studies. *Pharm Stat*. 2005;4:267-285.
35. Liang KY, Zeger SL. Longitudinal data analysis using generalized linear models. *Biometrika*. 1986;73:13-22.
36. Robins J, Rotnitzky A, Zhao LP. Analysis of semiparametric regression models for repeated outcomes in the presence of missing data. *J Am Stat Assoc*. 1995;90:106-121.
37. Collins LM, Schafer JL, Kam CM. A comparison of inclusive and restrictive missing-data strategies in modern missing-data procedures. *Psychol Methods*. 2001;6:330-351.
38. Papakostas GI, Petersen TJ, Iosifescu DV, Summergrad P, Sklarsky KG, Alpert JE, et al. Somatic symptoms as predictors of time to onset of response to fluoxetine in major depressive disorder. *J Clin Psychiatry*. 2004;65:543-546.
39. Jahn H, Schick M, Kiefer F, Kellner M, Yassouridis A, Wiedemann K. Metyrapone as additive treatment in major depression: a double-blind and placebo-controlled trial. *Arch Gen Psychiatry*. 2004;61: 1235-1244.
40. Stassen HH, Angst J, Delini-Stula A. Severity at baseline and onset of improvement in depression. Meta-analysis of imipramine and moclobemide versus placebo. *Eur Psychiatry*. 1994;9:129-136.
41. Stassen HH, Angst J, Delini-Stula A. Delayed onset of action of antidepressant drugs? Survey of results of Zurich meta-analyses. *Pharmacopsychiatry*. 1996;29:87-96.
42. Stassen HH, Angst J, Delini-Stula A. Fluoxetine versus moclobemide: cross-comparison between the time courses of improvement. *Pharmacopsychiatry*. 1999;32:56-60.
43. Artigas F. Limitations to enhancing the speed of onset of antidepressants—are rapid action antidepressants possible? *Hum Psychopharmacol*. 2001;16:29-36.
44. Katz MM, Koslow SH, Frazer A. Onset of antidepressant activity: reexamining the structure of depression and multiple actions of drugs. *Depress Anxiety*. 1996;4:257-267.
45. Thompson C. Onset of action of antidepressants: results of different analyses. *Hum Psychopharmacol*. 2002;17(Suppl 1):S27-S32.