



Published in final edited form as:

J Psychiatr Res. 2012 June ; 46(6): 708–715. doi:10.1016/j.jpsychires.2012.03.002.

Recovery and subsequent recurrence in patients with recurrent major depressive disorder

Boadie W. Dunlop^{a,*}, Peter Holland^b, Weihang Bao^c, Philip T. Ninan^c, and Martin B. Keller^d

^aEmory University School of Medicine, 1256 Briarcliff Road, Building A, 3rd Floor, Atlanta, GA 30322, USA

^bCharles E. Schmidt College of Medicine, Florida Atlantic University, 777 Glades Rd # 287 Boca Raton, FL 33431, USA

^cPfizer Inc, 500 Arcola Rd., Collegetown, PA 19426, USA

^dBrown University, 700 Butler Drive, Providence, RI 02906, USA

Abstract

In contrast to “remission” from an episode of major depressive disorder (MDD), for which there is general agreement in the literature, the optimal definition of “recovery” from MDD is uncertain. Previous definitions of recovery have used inconsistent thresholds for symptom severity and duration of wellness. To address the effects of duration and degree of recovery from an episode of MDD on recurrence risk, and the impact of maintenance antidepressant treatment on recurrence, we analyzed 258 patients from a randomized, double-blind study of outpatients with recurrent MDD. All patients had responded to 8½ months of venlafaxine extended release and were subsequently randomized to receive venlafaxine ER or placebo during 2 consecutive 12-month maintenance phases. Four definitions of recovery were used to evaluate recovery rates and time to recurrence: (1) 17-item Hamilton Depression Rating Scale (HAM-D₁₇) total score ≤ 3 with duration ≥ 120 days; (2) HAM-D₁₇ ≤ 3 with duration ≥ 56 days; (3) HAM-D₁₇ ≤ 7 with duration ≥ 120 days; and (4) HAM-D₁₇ ≤ 7 with duration ≥ 56 days. Recovery definitions using lower symptom severity and longer duration thresholds produced lower rates of recurrence. Patients on placebo were more likely to have a recurrence than patients on venlafaxine ER, with hazard ratio (HR) ranging from 2.5 among patients who recovered by the most relaxed criteria (definition 4), to 5.3 among patients who recovered by the most stringent criteria (definition 1). We conclude that protection against recurrence derives from the degree and duration of recovery, particularly for patients maintained on antidepressant medication.

Keywords

Antidepressants; Recurrence; Venlafaxine; Forecasting; Placebo

© 2012 Elsevier Ltd. All rights reserved.

*Corresponding author. Tel.: +1 404 727 8969; fax: +1 404 727 3700. bdunlop@emory.edu (B.W. Dunlop).

¹Use Zip 30306 for courier services.

Contributors

Drs. Dunlop and Holland led the manuscript development. Dr. Bao conducted the statistical analysis. Dr. Ninan managed the dataset used for the analysis. Dr. Keller designed the study. All authors contributed to the manuscript and approved its final version.

Conflicts of interest

1. Introduction

The goal of antidepressant treatment of major depressive disorder (MDD) is recovery from the episode. Recovery is thought to provide the best protection against the return of a new depressive episode (ie, recurrence), and is generally regarded as maintaining remission (usually defined as a rating scale score below a specific threshold) for a given period of time. However, specific time- and symptom level criteria used to define recovery vary considerably between studies. Enhancing the definition of recovery would directly impact clinical care, as it would inform decisions about whether to augment a patient's current treatment, and when discontinuation of antidepressant medication is appropriate.

The MacArthur Foundation task force proposed 3 different sets of operational criteria for recovery, based on different assessment scales and using different durations [i.e., Schedule for Affective Disorders and Schizophrenia (Spitzer et al., 1978) 2 symptoms for at least 8 weeks, 17-item Hamilton Rating Scale for Depression (HAM-D₁₇; Hamilton, 1960) 7 for at least 6 months, and Beck Depression Inventory (Beck et al., 1961) 8 for at least 4 months] (Frank et al., 1991). The definition used in the National Institute of Mental Health Collaborative Depression Study specified the presence of no symptoms or 1–2 symptoms to a mild degree for a minimum of 8 consecutive weeks (Keller et al., 1983; Solomon et al., 1997). The *Diagnostic and Statistical Manual of Mental of Mental Disorders* (Fourth Edition, Text Revision) (DSM-IV-TR) defines the end of a depressive episode as a period of at least 2 months during which full criteria for a major depressive episode are not met, although this relatively weak definition allows for significant ongoing symptomatology (American Psychiatric Association, 1994). Other studies of long-term outcomes defined recovery using a minimum duration of 3 months (Spijker et al., 2002; Yiend et al., 2009).

Recurrence of a depressive episode after recovery is thought to be related to both to the level of symptoms present during recovery, and the duration of the recovery period. The presence of residual symptoms is associated with an increased risk of recurrence and a shorter time to recurrence (Judd et al., 1998, 2000), whereas a longer duration of recovery is associated with a lower risk of recurrence (Solomon et al., 2000). Thus, the criteria used to define recovery may have implications when considering the impact of recovery on long-term outcomes. For example, in a recent analysis of outcomes during 10 years of follow-up of patients with MDD, recovery defined using a duration of 4–6 months was associated with a median time to subthreshold recurrence (no longer meeting criteria for recovery, but not meeting full MDD criteria) of 3 years, whereas a duration of 2 months was associated with subthreshold recurrence within 1.5 years for more than half of patients (Furukawa et al., 2008).

Maintenance treatment with antidepressants is effective in reducing rates of recurrence as well as increasing time to recurrence in patients with a history of recurrent depression (Lepine et al., 2004; Hochstrasser et al., 2001; Kocsis et al., 2007; Keller et al., 2007a,b; Hansen et al., 2008). However, very little placebo-controlled research has examined whether gradations of residual symptoms in patients meeting remission criteria (i.e., HAM-D₁₇ 7) or whether variability in the duration of sustained remission differentially impact rates of recurrence. It is also unknown whether antidepressant treatment provides differing levels of protection against recurrence among patients achieving different levels of recovery. Exploring these questions are important for understanding differential recurrence risks among “recovered” patients, and to determine the degree to which maintenance antidepressant treatment provides added benefit in preventing recurrences.

1.1. Objectives of the study

This analysis was conducted to assess rates of recovery during up to 2 years of maintenance treatment with venlafaxine extended release (ER) or placebo in patients with recurrent MDD, and to evaluate the effects of different definitions of recovery on time to and probability of recurrence. We hypothesized that a definition of recovery incorporating lower thresholds for symptom severity and a longer duration at that threshold would predict lower recurrence rates than the current standard definition of recovery. We also expected that the risk of depression recurrence between venlafaxine ER and placebo would be more evident in patients with more fragile recoveries (ie, short duration; higher symptom scores), as these patients may be most vulnerable to recurrence and therefore most in need of continued antidepressant treatment.

2. Methods

We conducted a post hoc analysis on the randomized sample of 258 patients from the Prevention of Recurrent Episodes of Depression with Venlafaxine for Two Years (PREVENT) trial (Kocsis et al., 2007; Keller et al., 2007a, b), a multiphase, double-blind trial of adult outpatients with recurrent MDD. The study was conducted in accord with the Declaration of Helsinki and its amendments. The study was reviewed and approved by the ethics review body responsible for each site, and all participants provided written informed consent prior to any study procedures being performed.

A schematic diagram of the PREVENT trial was previously published (Thase et al., 2011). In the PREVENT trial, patients were randomly assigned to 10-week double-blind acute treatment with either flexible-dose venlafaxine ER (75–300 mg/d) or fluoxetine (20–60 mg/d). Patients who met criteria for response (HAM-D₁₇ total score \leq 12 and \geq 50% reduction from acute phase baseline) at the end of acute treatment entered a 6-month continuation phase on the same double-blind medication. Responders at the end of the continuation phase were enrolled in the first of 2 consecutive double-blind 12-month maintenance phases (A and B). There were 258 patients in the venlafaxine ER group who were randomly assigned to venlafaxine ER or placebo and were evaluable for efficacy during maintenance phase A. Patients responding to venlafaxine ER at the end of maintenance phase A were again randomized to receive either venlafaxine ER or placebo during maintenance phase B. Non-relapsing placebo-treated patients from phase A continued on placebo in phase B. Data from both maintenance phases A and B will be used in the current analysis. Each participating clinical site for the PREVENT trial received approval to conduct the study from their respective institutional review board.

2.1. Patients

Eligible participants were outpatients \geq 18 years of age with recurrent MDD (defined as \geq 3 lifetime major depressive episodes, with \geq 2 episodes including the current episode occurring in the past 5 years, and with an interval \geq 2 months between the end of the previous episode and the beginning of the present episode). Episode onset and resolution were defined per Structured Clinical Interview for DSM-IV-TR criteria (First et al., 2007). Patients were required to have a HAM-D₁₇ total score \leq 20 at screening and \leq 18 at randomization and meet DSM-IV-TR criteria for major depressive episode for \geq 1 month prior to study entry.

Major exclusion criteria included failure to respond to an adequate trial of fluoxetine, venlafaxine, or venlafaxine ER during the current episode of MDD; known hypersensitivity, previous intolerance, or unsuccessful treatment with venlafaxine or fluoxetine; previous treatment resistance, defined as having failed in the past 3 years: (a) \geq 3 previous adequate trials of \geq 2 classes of antidepressants; (b) electroconvulsive therapy; or (c) 2 adequate trials

of psychotherapy; history/presence of bipolar disorder or eating disorder (if not remitted for 5 years), or significant axis II disorders; or a primary diagnosis of panic disorder, obsessive compulsive disorder, generalized anxiety disorder, social phobia, or posttraumatic stress disorder within 6 months prior to screening.

At acute phase baseline, 821 patients were randomized to receive venlafaxine ER, 530 of whom entered into the continuation phase. Patients included in the current analyses had all responded to the acute 10-week treatment course of flexibly dosed venlafaxine ER, and had maintained their response during the 6-month continuation phase.

2.2. Study assessments

Patients visited the study sites weekly or biweekly during the acute treatment phase, and monthly throughout the 6-month continuation phase and both 1-year maintenance phases. The HAM-D₁₇ was performed at each visit throughout the study. Vital signs and adverse events (AEs) were collected at each visit, and laboratory evaluations were performed at screening, at the end of the continuation phase, and at the last visit of each maintenance phase. Detailed information on study protocol and assessments has been previously published (Keller et al., 2007b).

2.3. Outcomes

2.3.1. Definitions of recovery—To evaluate recovery, symptom severity (based on HAM-D₁₇ total score) and the duration of remission were considered. Two thresholds of symptom severity were used. Consistent with published literature for remission (Frank et al., 1991), a HAM-D₁₇ total score ≤ 7 was used as one threshold. A second threshold, employing a more stringent criterion of HAM-D₁₇ ≤ 3 , was used to distinguish between patients who had mild residual symptoms versus those who were essentially asymptomatic (Judd et al., 1998). The basis for this cut-off was supported by studies conducted by 2 independent groups. Zimmerman et al. (2004) found that the mean HAM-D₁₇ score among healthy controls is 3.2 (95% confidence interval: 3.0–3.4), and Zimmerman et al. (2005) demonstrated that threshold definitions of remission <7 on the HAM-D₁₇ correlate best with self-reported psychosocial impairment and quality of life. Separately, Furukawa et al. (2007) reported that an end-of-treatment CGI-Severity (CGI-S) score of 1 (“Not at all ill”) correlated with a HAM-D₁₇ score ≤ 3 , whereas a HAM-D₁₇ score between 4 and 7 corresponded to a CGI-S score of 2 (“Borderline ill”). Two recovery duration criteria (8 weeks [56 days] and 4 months [120 days]) were employed based on previously used duration definitions (Frank et al., 1991; Keller et al., 1983; Furukawa et al., 2008) and those recommended by the ACNP Task Force (Rush et al., 2006). For the purposes of this analysis, 4 definitions of recovery were used: (1) HAM-D₁₇ ≤ 3 with duration ≥ 120 days, (2) HAM-D₁₇ ≤ 3 with duration ≥ 56 days, (3) HAM-D₁₇ ≤ 7 with duration ≥ 120 days, and (4) HAM-D₁₇ ≤ 7 with duration ≥ 56 days.

2.3.2. Effect of recovery on recurrence—Analyses were performed to evaluate the impact of the 4 recovery definitions on: (1) the probability of achieving recovery versus non-recovery; (2) the probability of and time to recurrence; and (3) the effect of treatment (venlafaxine ER vs placebo) on recurrence risks. Recurrence was defined as a HAM-D₁₇ score >12 and a $<50\%$ decrease from acute phase baseline at 2 consecutive visits or at the last valid visit before discontinuation.¹⁷

2.4. Statistical analysis

2.4.1. Rates of recovery—Statistical analyses were based on the 258 patients (the “efficacy evaluable sample”) who were randomized to venlafaxine ER or placebo at the end of the continuation phase and who had efficacy evaluable data. This sample included all

randomized patients in the intent-to-treat (ITT) population of maintenance phase A (ie, all patients who took 1 dose of study medication and had 1 post-randomization HAM-D₁₇ assessment) except those who were affected by a drug-dispensing issue, described previously (Kocsis et al., 2007). Among the subset of patients ($n = 83$) who were randomized to venlafaxine ER or placebo group in maintenance phase B, rates of recovery were also calculated separately based on the ITT population. In addition, analyses were conducted for both maintenance phases combined. For the combined maintenance phases, if a patient was treated with venlafaxine ER and did not meet the definition requirement for recovery at end of maintenance A and was randomly assigned to placebo in maintenance B, the patient was considered not recovered and data from maintenance B were excluded.

Separately under each definition of recovery, the proportion of patients achieving recovery was compared between the treatment groups in a logistic regression analysis. Nominal (unadjusted) P values (2-sided) will be reported for all statistical comparisons.

2.4.2. Effect of recovery on recurrence—We performed 2 sets of analyses on the effect of recovery on recurrence. The first, (“inclusive analysis”), assessed the treatment effect on time to recurrence during the combined maintenance phases for each definition of recovery using Cox proportional hazards models. For this set of analyses, time to recurrence was measured from the beginning of recovery starting in maintenance phase A; therefore, only patients who met criteria for recovery during the maintenance phases were included in these analyses. Kapla–Meier plots were produced to graphically represent recurrence risks; product limit estimates were used to estimate recurrence proportions. This analytic approach is similar to those used in previous studies examining recovery in MDD, in that recovered patients are analyzed as inclusive groups; ie, those who met criteria for more stringent definitions of recovery are also included in the patient groups meeting criteria for less stringent definitions. For example, patients meeting recovery criterion 1 (the most stringent) are also included as meeting recovery criteria 2, 3, and 4.

In the second approach (“exclusive analysis”), each patient was counted only once, based on the most stringent definition of recovery they achieved. This analysis included all patients, separated them into 4 distinct groups, and aimed to better discriminate the effects of degree of recovery on recurrence risk. For the purposes of this analysis, time to recurrence was measured from the beginning of maintenance phase A. Patients were divided into 4 recovery groups:

- Group A: Patients recovered according to definition 1, the most stringent definition, ie, HAM-D₁₇ ≤ 3 with duration ≥ 120 days.
- Group B: Patients recovered according to the definition 2 or 3, but not 1, ie, HAM-D₁₇ ≤ 3 with duration ≥ 56 but <120 days, or HAM-D₁₇ ≤ 7 but >3 , with duration ≥ 120 days.
- Group C: Patients recovered according to definition 4, the most relaxed definition, ie, HAM-D₁₇ ≤ 7 but >3 , with duration ≥ 56 but <120 days.
- Group D: Patients who responded during the acute and continuation phases and entered the maintenance phases, but who never met criteria for any of the 4 definitions of recovery.

To examine the effect of recovery on time to recurrence, Cox proportional hazards models were used with terms for patient recovery group and treatment. The patient recovery group by treatment interaction was also examined in a separate model. Because more patients treated with venlafaxine maintenance treatment achieved recovery or recovered more completely, analysis of treatment effect within each recovery category is subject to patient

selection bias, underestimating the treatment benefit. Therefore, these Cox proportional analyses are not intended to quantify treatment benefit.

2.4.3. Discontinuations—Rates of discontinuation were summarized for patients meeting each definition of recovery (ie, definition 1, 2, 3, or 4); total discontinuations, discontinuations due to unsatisfactory response, and discontinuations due to AEs were included.

3. Results

3.1. Patients and dosage

The efficacy evaluable sample (for maintenance A and the combined maintenance phases) included 258 patients (placebo, $n = 129$; venlafaxine ER, $n = 129$); 83 of these patients were included in the maintenance B ITT population (placebo, $n = 40$; venlafaxine ER, $n = 43$). In both phases, demographic and baseline clinical characteristics were similar for venlafaxine ER and placebo groups; details of this sample have been previously reported (Kocsis et al., 2007; Keller et al., 2007a). Briefly, the sample had a mean age of 42.3 years, was 68% female and 85% white, had a mean duration of current episode of 6.7 months, and a mean HAM-D₁₇ score at entry into maintenance phase A of 4.6 (Kocsis et al., 2007). The mean venlafaxine ER daily dose was 220.8 ± 71.8 mg during maintenance A (Kocsis et al., 2007) and 213.5 ± 75.2 mg during maintenance B (Keller et al., 2007a).

3.2. Outcomes

3.2.1. Rates of recovery—Fig. 1/Table 1 show the percentage of patients in recovery by treatment group, criterion, and study phase. In the individual and combined maintenance phases, venlafaxine ER was consistently associated with higher rates of recovery compared with placebo regardless of criteria for recovery, with odds ratios ranging from 2.41 among patients who recovered by the most relaxed criteria (definition 4), to 1.57 among patients who recovered by the most stringent criteria (definition 1). Statistical significance was reached for all definitions except for definition 1 in both maintenance phases and definition 1 in maintenance phase B. Only 35.8% (58/162) patients who met the least strict definition (4) for recovery met criteria for the strictest definition (1).

3.2.2. Recurrence risk—The results of the inclusive analysis showing the proportion of subjects who had a recurrence during the 2-year maintenance period is presented in Table 2 by each definition of recovery. Patients on placebo were more likely to have a recurrence than patients on venlafaxine in the combined maintenance phases. The hazard ratio (HR) ranged from 2.5 among patients who recovered by the most relaxed criteria (definition 4), to 5.3 among patients who recovered by the most stringent criteria (definition 1). These results indicate that venlafaxine provides substantial benefits in maintaining wellness beyond those provided by placebo, especially among patients who achieve very low symptom levels for a sustained period. Although definition 1 produced the largest HR, the smaller number of patients meeting this recovery definition resulted in a failure to reach statistical significance. Statistical significance was achieved for recovery definitions 2, 3, and 4. Kaplan–Meier curves in Fig. 2 illustrate that recovered patients on venlafaxine maintenance treatment were more likely to remain recurrence-free, regardless of definition used for recovery.

Table 3 demonstrates the effects on recurrence from the exclusive analysis, in which patients are grouped by level of recovery, with each patient counted only once. Of the 258 patients who had responded by the end of continuation, 162 (62.8%) achieved at least one of the recovery definitions. Patients who achieved any recovery, regardless of criteria used, had significantly less risk of recurrence when compared to patients who did not achieve recovery

(group D). Furthermore, compared with patients who never met criteria for recovery (group D), patients meeting the most relaxed criteria (group C) reduced their risk by 86.6%. Patients who achieved most stringent criteria (group A) reduced their risk significantly more (98%, $P=0.0027$). Because more patients treated with venlafaxine achieved recovery or recovered more completely (Table 1), the direct comparison of recurrence risk between venlafaxine and placebo within each recovery group is subject to patient selection bias, underestimating the treatment benefit. Nevertheless, as shown in Table 3, the proportion of venlafaxine-treated patients having a recurrence in groups A and B is approximately one-quarter the rate of placebo-treated patients. In contrast, the proportions are roughly equal for recurrence risks in groups C and D. The overall treatment by recovery group interaction was not significant, indicating that the finding of more complete recovery led to less recurrence held true for patients who continued venlafaxine treatment and for those who received placebo ($P=0.24$).

3.2.3. Discontinuations—Discontinuation rates for the individual maintenance phases are presented in Table 4. In maintenance phase A, overall discontinuations across the definitions of recovery were generally greater with placebo (range 42–51%) than with venlafaxine ER (range 22–32%). Rates of discontinuations due to unsatisfactory response also were lower in the venlafaxine ER groups. In addition, discontinuations due to unsatisfactory response in both the venlafaxine ER and placebo groups were lower among patients who met definitions of recovery requiring 120 days (venlafaxine, 0–1%; placebo, 10%) compared with the corresponding definition using 56 days (venlafaxine, 3–7%; placebo 15%). Discontinuations due to AEs were slightly greater in the venlafaxine ER groups (2–3%) compared with the placebo groups (0–2%); however, there was no consistent pattern with respect to differences among the recovery subgroups. In maintenance phase B, discontinuations due to unsatisfactory response were generally greater with placebo and discontinuations due to AEs were generally greater with venlafaxine ER, although the small sample sizes limit interpretation of the results from this phase.

4. Discussion

All 4 definitions of recovery predicted a stable long-term course for most patients. Rates of recurrence were significantly lower for patients meeting any definition of recovery (exclusively defined recovery groups A–C), than for those achieving response short of recovery (group D). Better long-term outcomes were observed using stricter recovery definitions than the currently established definition of a HAM-D₁₇ 7, for 2 months' duration. Only a minority (36%) of the patients who met the established definition for recovery (definition 4) also met the criteria for the strictest definition (definition 1).

Contrary to our expectation, greater protective effects of maintenance antidepressant versus placebo occurred among the patients achieving the stricter definitions of recovery, based on the HRs using the inclusive analysis (definitions 1, 2, and 3 vs 4). This finding was supported by our exclusive analysis, in which patients with the most fragile recoveries (group C), and those who did not achieve recovery (group D), had no difference in recurrence rates for those maintained on venlafaxine versus placebo, whereas strong recurrence prevention effects of venlafaxine were observed in the more stringent recovery groups (groups A and B). Thus, our data suggest that the protective effect of medication is not a “step-function,” in which it is no longer of value once a certain level of recovery of achieved. Rather, these results indicate that the benefit of maintenance medication to prevent recurrence is a continuous function across symptom level and duration of wellness. Taken together, these findings suggest that the well-replicated findings of the protective effects of maintenance antidepressant treatment over placebo are driven by those patients with the most complete response to medication.

Patients who barely meet standard recovery criteria appear to contribute little to drug-placebo separation. However, as demonstrated by the lack of statistical significance from the overall treatment-by-group interaction in the exclusive analysis, using the inclusive standard definition may be necessary to obtain the statistical power needed to reach statistically significant results. Thus, for purposes of clinical trial design, the standard definition allows for the greatest ability to discriminate drug versus placebo differences, even though the effect size is smaller with this definition. Whether incorporation of self-reported outcomes may improve signal detection for recurrence studies warrants further investigation (Dunlop et al., 2011).

A potential explanation for why the protective effect of venlafaxine was smaller for the less strict definitions of recurrence can be seen from the numbers of patients achieving each definition of recovery (Table 1). The number of venlafaxine-treated patients was greater than the number of placebo-treated patients for each definition, with a correspondingly larger number of placebo-treated patients remaining in the non-recovered group (group D). This difference implies that venlafaxine treatment “pushed” many harder-to-treat patients (ie, those who would not recover with placebo) into recovery. These harder-to-treat patients may have had clinical or biological characteristics that made a recurrence more likely, producing a bias toward recurrence in the venlafaxine-treated “recovered” patients. If so, our analyses would underestimate the protective effect of venlafaxine versus placebo in maintenance treatment. This potential bias applies not only to the PREVENT dataset, but also to all similarly designed double-blind discontinuation studies for MDD. To provide an unbiased estimate of the protective effect in future studies, patients will have to be randomized to receive venlafaxine or placebo maintenance treatment at the time they achieve recovery.

Ten to 15% of patients on placebo who met recovery criteria discontinued prematurely from the study due to unsatisfactory response. These percentages are considerably higher than the discontinuation rate of 0% to 6.5% for patients maintained on venlafaxine. Patients who met criteria for recovery who later terminate the study due to unsatisfactory response are likely experiencing a return of significant depressive symptoms short of full recurrence criteria, requiring a modification of treatment. Thus, our reported effects of venlafaxine ER over placebo in preventing depressive recurrence may underestimate the actual protective effects, as breakthrough depressive symptoms leading to termination were not included in our primary outcome, but were more common in the placebo-treated group.

The major strength of this analysis is the longitudinal documentation of maintenance treatment outcomes on a large number of subjects with established recurrent MDD. However, the generalizability of the results to everyday clinical practice may be limited by several factors. First, this was a post hoc analysis of data from a long-term clinical trial. Patients enrolled did not have serious or unstable secondary medical conditions and thus may not be representative of typical depressed outpatients in clinical practice. Furthermore, the PREVENT trial enrolled only patients with a history of recurrent depression, so the results may not generalize to patients with only 1 or 2 lifetime episodes of depression. The study design, although not unusual for maintenance phase efficacy trials, introduced a selection bias in that only those patients who responded to treatment during the acute phase and did not relapse during the continuation phase entered the maintenance phase.

Another important limitation to this post hoc analysis is that, due to the study design, we were unable to randomize patients at the time that they actually achieved a specified recovery definition. Thus, these results should be recognized as preliminary conclusions warranting future replication in trials specifically designed to assess recovery. Better estimates of the protective effects of medication require prospective, randomized double-

blind, placebo-controlled discontinuation trials, in which patients are randomized at the time they meet pre-specified recovery definitions. The results from our analysis should inform the design and power calculations necessary for such trials.

Although among recovered patients the relative risk of recurrence is significantly greater among patients switched to placebo compared with those maintained on antidepressant medication, most recovered patients did not experience a recurrence during the 2 years of follow-up, regardless of treatment. Greater degree and longer duration of wellness provide additional protection against recurrence, and the gains from antidepressant treatment may require consolidation to have lasting effect. The decision to maintain antidepressant medication after achieving recovery should also consider patient-specific factors, such as previous suicidality and degree of functional impairment associated with depressive episodes.

Acknowledgments

This study and the statistical analysis for this manuscript were sponsored by Wyeth, which was acquired by Pfizer in October 2009. Medical writing and editing support for this manuscript was provided by Patricia Bakos, MS, RD, Sherri Jones, PharmD, and Jennifer Karpinski, BA, all of Embryon, LLC, A division of Advanced Health Media, LLC, and was funded by Wyeth.

Role of the funding source

This study was funded by Wyeth Research (now Pfizer).

This study and the statistical analysis for this manuscript were sponsored by Wyeth, which was acquired by Pfizer in October 2009. Drs. Bao and Ninan are employees of Pfizer. Drs. Dunlop, Holland and Keller have performed consulting for Wyeth in the past. In addition, in the past 3 years, Dr. Dunlop has served as a consultant to BMS, and received research support from AstraZeneca, BMS, Forest, GSK, Pfizer, Transcept and Wyeth. Dr. Holland has received research support from Forest Research Institute, Novartis, Takeda, Cyberonics, Eli Lilly and Genentech. He is also a consultant to Pfizer. In the past three years, Dr. Keller has consulted to Medtronic, Sierra Neuropharmaceuticals and CENEREX; he is on the scientific advisory board of CENEREX and has received research grants from Wyeth and Pfizer.

References

- American Psychiatric Association. American Psychiatric Association. 4th ed.. 1994. Diagnostic and statistical manual of mental disorders DSM-IV.
- Beck AT, Ward CH, Mendelson M, Mock J, Erbaugh J. An inventory for measuring depression. *Archives of General Psychiatry*. 1961; 4:561–571. [PubMed: 13688369]
- Dunlop BW, Li T, Kornstein SG, Friedman ES, Rothschild AJ, Pederson R, et al. Concordance between clinician and patient ratings as predictors of response, remission, and recurrence in major depressive disorder. *Journal of Psychiatric Research*. 2011; 45:96–103. [PubMed: 20537348]
- First, MB.; Gibbon, M.; Spitzer, RL., et al. Structured clinical interview for DSM-IV-TR axis I disorders, clinical trials version (SCID-CT). New York, NY: Biometrics Research, New York State Psychiatric Institute; 2007.
- Frank E, Prien RF, Jarrett RB, Keller MB, Kupfer DJ, Lavori PW, et al. Conceptualization and rationale for consensus definitions of terms in major depressive disorder: remission, recovery, relapse, and recurrence. *Archives of General Psychiatry*. 1991; 48:851–855. [PubMed: 1929776]
- Furukawa TA, Akechi T, Azuma H, Okuyama T, Higuchi T. Evidence-based guidelines for interpretation of the Hamilton rating scale for depression. *Journal of Clinical Psychopharmacology*. 2007; 27:531–534. [PubMed: 17873700]
- Furukawa TA, Fujita A, Harai H, Yoshimura R, Kitamura T, Takahashi K. Definitions of recovery and outcomes of major depression: results from a 10-year follow-up. *Acta Psychiatrica Scandinavica*. 2008; 117:35–40. [PubMed: 17986318]
- Hamilton M. A rating scale for depression. *Journal of Neurology Neurosurgery and Psychiatry*. 1960; 23:56–62.

- Hansen R, Gaynes B, Thieda P, Gartlehner G, Deveaugh-Geiss A, Krebs E, et al. Meta-analysis of major depressive disorder relapse and recurrence with second-generation antidepressants. *Psychiatric Services*. 2008; 59:1121–1130. [PubMed: 18832497]
- Hochstrasser B, Isaksen PM, Koponen H, Lauritzen L, Mahnert FA, Rouillon F, et al. Prophylactic effect of citalopram in unipolar, recurrent depression: placebo-controlled study of maintenance therapy. *British Journal of Psychiatry*. 2001; 178:304–310. [PubMed: 11282808]
- Judd LL, Paulus LL, Schettler PJ, Akiskal HS, Endicott J, Leon AC, et al. Does incomplete recovery from first lifetime major depressive episode herald a chronic course of illness? *American Journal of Psychiatry*. 2000; 157:1501–1504. [PubMed: 10964869]
- Judd LL, Akiskal HS, Maser JD, Zeller PJ, Endicott J, Coryell W, et al. Major depressive disorder: a prospective study of residual subthreshold depressive symptoms as predictor of rapid relapse. *Journal of Affective Disorders*. 1998; 50:97–108. [PubMed: 9858069]
- Keller MB, Lavori PW, Lewis CE, Klerman GL. Predictors of relapse in major depressive disorder. *Journal of the American Medical Association*. 1983; 250:3299–3304. [PubMed: 6645026]
- Keller MB, Trivedi MH, Thase ME, Shelton RC, Kornstein SG, Nemeroff CB, et al. The prevention of recurrent episodes of depression with venlafaxine for two years (PREVENT) study: outcomes from the 2-year and combined maintenance phases. *Journal of Clinical Psychiatry*. 2007a; 68:1246–1256. [PubMed: 17854250]
- Keller MB, Trivedi MH, Thase ME, Shelton RC, Kornstein SG, Nemeroff CB, et al. The prevention of recurrent episodes of depression with venlafaxine for two years (PREVENT) study: outcomes from the acute and continuation phases. *Biological Psychiatry*. 2007b; 62:1371–1379. [PubMed: 17825800]
- Kocsis JH, Thase ME, Trivedi MH, Shelton RC, Kornstein SG, Nemeroff CB, et al. Prevention of recurrent episodes of depression with venlafaxine ER in a 1-year maintenance phase from the PREVENT study. *Journal of Clinical Psychiatry*. 2007; 68:1014–1023. [PubMed: 17685736]
- Lepine JP, Caillard V, Bisslerbe JC, Troy S, Hotton JM, Boyer P. A randomized, placebo-controlled trial of sertraline for prophylactic treatment of highly recurrent major depressive disorder. *American Journal of Psychiatry*. 2004; 161:836–842. [PubMed: 15121648]
- Rush AJ, Kraemer HC, Sackeim HA, Fava M, Trivedi MH, Frank E, et al. ACNP task force. Report by the ACNP task force on response and remission in major depressive disorder. *Neuropsychopharmacology*. 2006; 31:1841–1853. [PubMed: 16794566]
- Solomon DA, Keller MB, Leon AC, Mueller TI, Shea MT, Warshaw M, et al. Recovery from major depression: a 10-year prospective follow-up across multiple episodes. *Archives of General Psychiatry*. 1997; 54:1001–1006. [PubMed: 9366656]
- Solomon DA, Keller MB, Leon AC, Mueller TI, Lavori PW, Shea MT, et al. Multiple recurrences of major depressive disorder. *American Journal of Psychiatry*. 2000; 157:229–233. [PubMed: 10671391]
- Spijker J, De GR, Bijl RV, Beekman AT, Ormel J, Nolen WA. Duration of major depressive episodes in the general population: results from the Netherlands mental health survey and incidence study (NEMESIS). *British Journal of Psychiatry*. 2002; 181:208–213. [PubMed: 12204924]
- Spitzer RL, Endicott J, Robins E. Research diagnostic criteria: rationale and reliability. *Archives of General Psychiatry*. 1978; 35:773–782. [PubMed: 655775]
- Thase ME, Gelenberg A, Kornstein SG, Kocsis JH, Trivedi MH, Ninan P, et al. Comparing venlafaxine extended release and fluoxetine for preventing the recurrence of major depression: results from the PREVENT study. *Journal of Psychiatric Research*. 2011; 45:412–420. [PubMed: 20801464]
- Yiend J, Paykel E, Merritt R, Lester K, Doll H, Burns T. Long term outcome of primary care depression. *Journal of Affective Disorders*. 2009; 118:79–86. [PubMed: 19246103]
- Zimmerman M, Chelminski I, Posternak M. A review of studies of the Hamilton depression rating scale in healthy controls: implications for the definition of remission in treatment studies of depression. *Journal of Nervous and Mental Diseases*. 2004; 192:595–601.
- Zimmerman M, Posternak MA, Chelminski I. Is the cutoff to define remission on the Hamilton rating scale for depression too high? *Journal of Nervous and Mental Diseases*. 2005; 193:170–175.

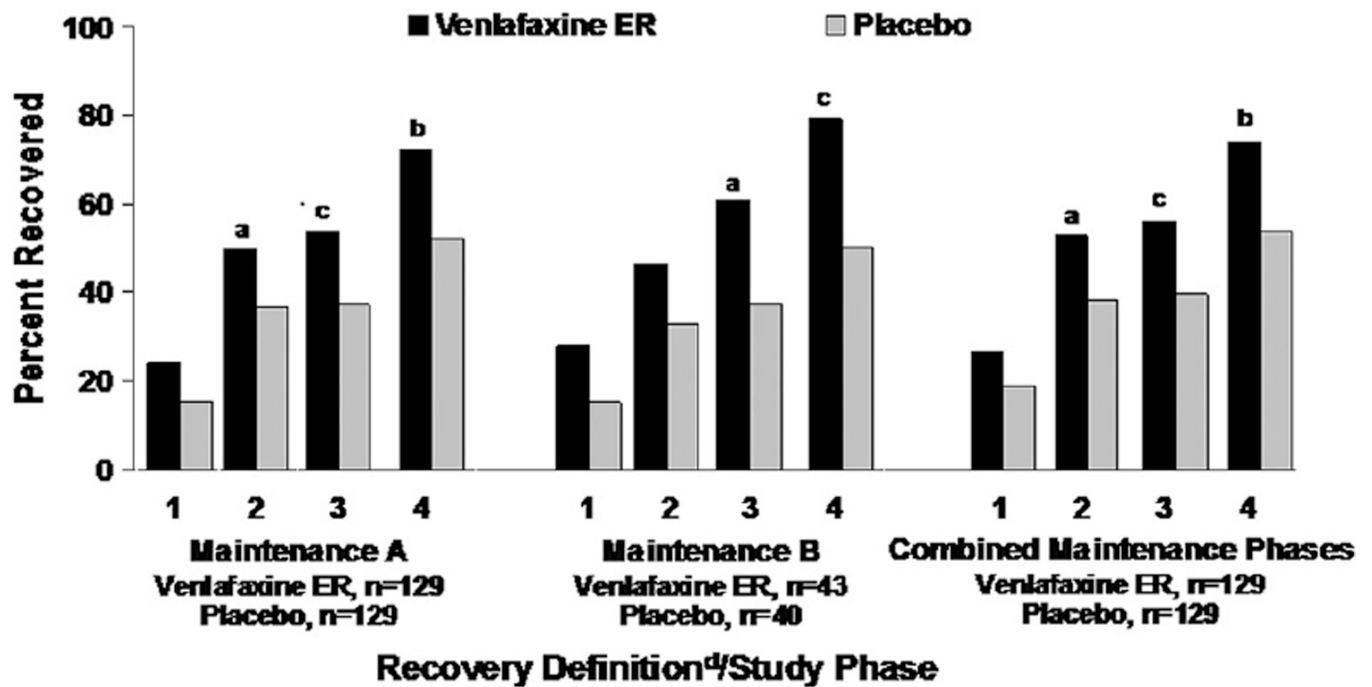


Fig. 1.

Percent of patients achieving recovery by criteria and study phase. Abbreviations: ER = extended release; HAM-D₁₇, 17-item Hamilton rating scale for depression. ^a $P < 0.05$ compared with placebo. ^b $P < 0.001$ compared with placebo. ^c $P < 0.01$ compared with placebo. ^dRecovery definitions: (1) HAM-D₁₇ total score ≤ 3 with duration ≥ 120 days. (2) HAM-D₁₇ total score ≤ 3 with duration ≥ 56 days. (3) HAM-D₁₇ total score ≤ 7 with duration ≥ 120 days. (4) HAM-D₁₇ total score ≤ 7 with duration ≥ 56 days.

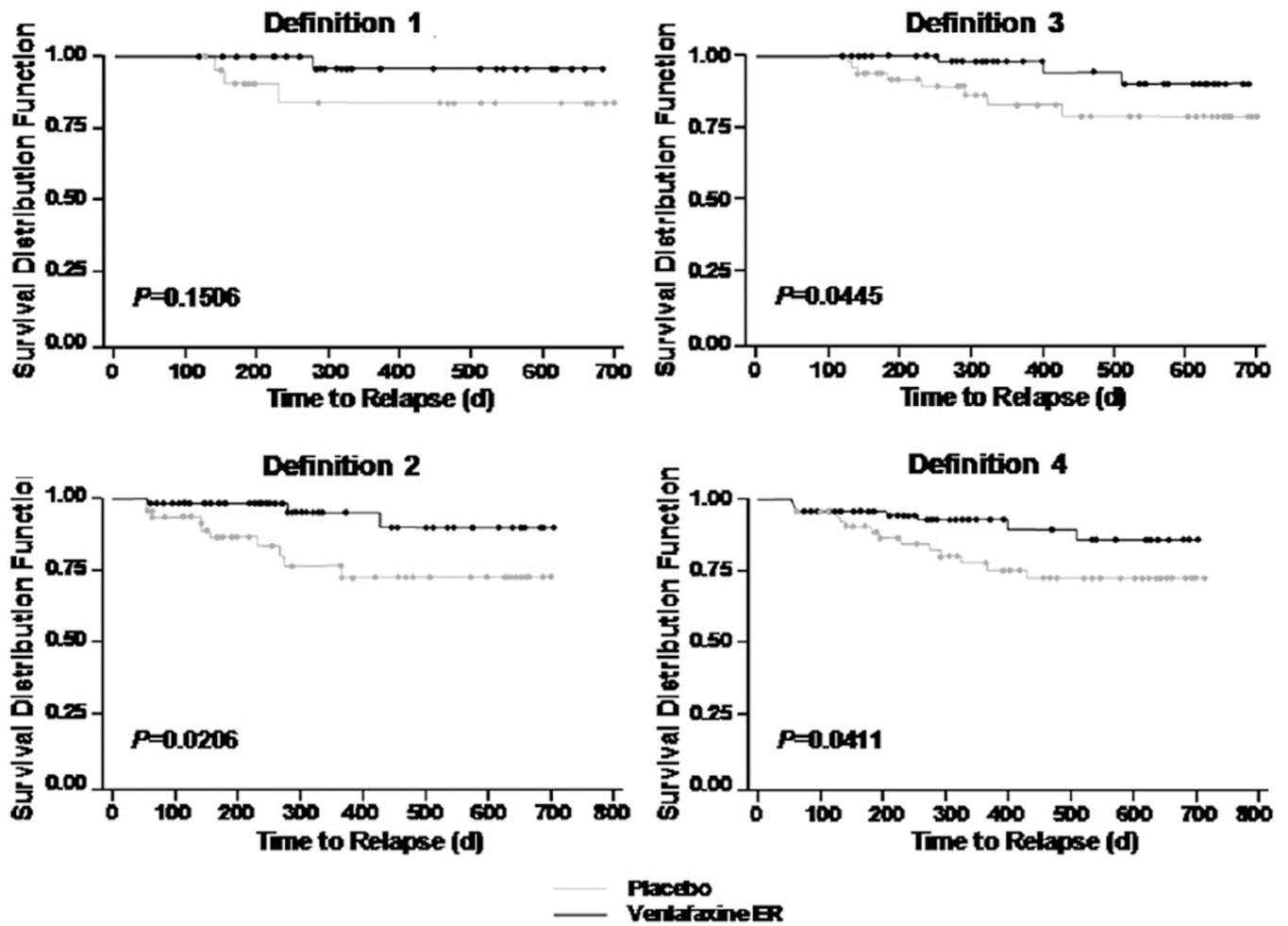


Fig. 2. Kaplan–Meier curves for probability of no recurrence by definition of recovery. Circles indicate censored values.

Table 1

Subjects achieving recovery by definition in venlafaxine ER vs placebo groups during the combined maintenance phase.

Venlafaxine/placebo analysis^a				
Efficacy evaluable	Venlafaxine ER n (%)	Placebo n (%)	P value	OR (95% CI)
Total <i>N</i>	129	129		
Recovery definition 1	34 (26.4)	24 (18.6)	0.1792	1.566 (0.867, 2.829)
Recovery definition 2	68 (52.7)	49 (38.0)	0.0242	1.820 (1.108, 2.988)
Recovery definition 3	72 (55.8)	51 (39.5)	0.0125	1.932 (1.177, 3.170)
Recovery definition 4	94 (72.9)	68 (52.7)	0.0012	2.409 (1.433, 4.051)

Abbreviations: CI, confidence interval; ER, extended release; HAM-D₁₇, 17-item Hamilton rating scale for depression; OR, odds ratio.

Recovery definitions:(1) HAM-D₁₇ total score ≤ 3 with duration ≥ 120 days.(2) HAM-D₁₇ total score ≤ 3 with duration ≥ 56 days.(3) HAM-D₁₇ total score ≤ 7 with duration ≥ 120 days.(4) HAM-D₁₇ total score ≤ 7 with duration ≥ 56 days.

^aAnalysis of data from maintenance A and B combined; patients meeting recovery during maintenance phases.

Table 2

Hazard ratios for timetorecurrence during the combined maintenance phase for the 4 definitions of recovery (“inclusive” analysis).

Recovery definition	Placebo vs venlafaxine ER ^a			
	1	2	3	4
Total events (%)	4/54 (7.4)	13/117 (11.1)	11/123 (8.9)	22/162 (13.6)
Placebo events (%)	3/24 (12.5)	10/49 (20.4)	8/51 (15.7)	14/68 (20.6)
Venlafaxine events (%)	1/34 (2.9)	3/68 (4.4)	3/72 (4.2)	8/94 (8.5)
HR	5.28	4.61	3.92	2.48
95% CI	(0.55, 51.00)	(1.26, 16.82)	(1.03, 14.83)	(1.04, 5.94)
<i>p</i> Value	0.1506	0.0206	0.0445	0.0411

Abbreviations: ER, extended release; HAM-D₁₇, 17-item Hamilton rating scale for depression; HR, hazard ratio.

Recovery definitions: (1) HAM-D₁₇ total score ≤ 3 with duration ≥ 120 days. (2) HAM-D₁₇ total score ≤ 3 with duration ≥ 56 days. (3) HAM-D₁₇ total score ≤ 7 with duration ≥ 120 days. (4) HAM-D₁₇ total score ≤ 7 with duration ≥ 56 days.

^aAnalysis of data from efficacy evaluable population from maintenance A and B combined; patients meeting recovery during maintenance phases.

Table 3

Hazard ratios comparing time to recurrence by strictest level of recovery achieved (“exclusive” analysis).

Recurrence analysis^a					
Recovery group^b	A	B	C	D	
Patient recurrence, n/N (%)					
Total	4/58 (6.9)	11/79 (13.9)	7/25 (28.0)	53/96 (55.2)	
Placebo	3/24 (12.5)	8/32 (25.0)	3/12 (25.0)	32/61 (52.5)	
Venlafaxine ER	1/34 (2.9)	3/47 (6.4)	4/13 (30.8)	21/35 (60.0)	
Between-group comparisons (placebo and venlafaxine or patients combined)					
HR vs group D	0.020	0.046	0.134	–	
<i>P</i> value	<0.0001	<0.0001	<0.0001	–	
HR vs group C	0.150	0.347	–	–	
<i>P</i> value	0.0027	0.0297	–	–	
HR vs group B	0.432	–	–	–	
<i>P</i> value	0.1517	–	–	–	

Abbreviations: ER, extended release; HAM-D17, 17-item Hamilton rating scale for depression; HR, hazard ratio.

^aRecurrence is defined as HAM-D17 total score >12, <50% reduction in the HAM-D17 total as compared with baseline (acute phase), *Diagnostic and Statistical Manual of Mental Disorders* (Fourth Edition, Text Revision) criteria met for major depression, and independent confirmation of criteria by a senior investigator.

^bPatient recovery group definitions:(A)HAM-D17 total score 3 with duration 120 days;(B)HAM-D17 total score 7 with duration 120 days, or HAM-D17 total score 3 with duration 56 days; and excluding patients from group A.(C)HAM-D17 total score 7 with duration 56 days; and excluding patients from groups A and B.(D)Patients who did not recover; ie patients not in groups A, B, or C.

Table 4

Discontinuations.

Maintenance A ^a	Total N		Total discontinued (%)		Unsatisfactory response (%)		Adverse event (%)	
	Venlafaxine ER	Placebo	Venlafaxine ER	Placebo	Venlafaxine ER	Placebo	Venlafaxine ER	Placebo
<i>56 days</i>								
HAM-D ₁₇ 3 (definition 2)	64	47	29.7	51.1	3.1	14.9	1.6	2.1
HAM-D ₁₇ 7 (definition 4)	93	67	32.3	49.3	6.5	14.9	3.2	1.5
<i>120 days</i>								
HAM-D ₁₇ 3 (definition 1)	31	20	23.6	50.0	0	10.0	3.2	0
HAM-D ₁₇ 7 (definition 3)	69	48	21.7	41.7	1.4	10.4	2.9	2.1
Maintenance B ^b	Venlafaxine ER	Placebo	Venlafaxine ER	Placebo	Venlafaxine ER	Placebo	Venlafaxine ER	Placebo
<i>56 days</i>								
HAM-D ₁₇ 3 (definition 2)	20	13	15.0	23.1	0	15.4	5.0	0
HAM-D ₁₇ 7 (definition 4)	34	20	14.7	30.0	2.9	25.0	2.9	0
<i>120 days</i>								
HAM-D ₁₇ 3 (definition 1)	12	6	8.3	0	0	0	0	0
HAM-D ₁₇ 7 (definition 3)	26	15	11.5	6.7	0	6.7	3.8	0

Abbreviations: ER, extended release; HAM-D₁₇, 17-item Hamilton rating scale for depression.

Recovery definitions:(1)HAM-D₁₇ total score 3 with duration 120 days.(2)HAM-D₁₇ total score 3 with duration 56days.(3)HAM-D₁₇ total score 7 with duration 120 days.(4) HAM-D₁₇ total score 7 with duration 56 days.

^aEfficacy-evaluable population.

^bIntent-to-treat population.