PATTERN ANALYSIS OF RESPONSE TO ACUTE FLUOXETINE TREATMENT IN THE PREDICTION OF RELAPSE IN CHILDREN AND ADOLESCENTS WITH MAJOR DEPRESSIVE DISORDER

APPROVED BY SUPERVISORY COMMITTEE

Carroll W. Hughes, Ph. D.

Graham J. Emslie, M.D.

Beth D. Kennard, Psy.D.

Thomas J. Carmody, Ph.D.

Deanna S. Liss, Ph.D.

to Paul

PATTERN ANALYSIS OF RESPONSE TO ACUTE FLUOXETINE TREATMENT IN THE PREDICTION OF RELAPSE IN CHILDREN AND ADOLESCENTS WITH MAJOR DEPRESSIVE DISORDER

by

ANN STEVENS AIRY EGGERTSEN

DISSERTATION

Presented to the Faculty of the Graduate School of Biomedical Sciences

The University of Texas Southwestern Medical Center at Dallas

In Partial Fulfillment of the Requirements

For the Degree of

DOCTOR OF PHILOSOPHY

The University of Texas Southwestern Medical Center at Dallas

Dallas, Texas

August 2008

ACKNOWLEDGEMENTS

There are several people whose guidance and support were instrumental to my successful completion of this doctoral program and dissertation. First, I want to express my appreciation to the members of my committee who provided valuable assistance throughout my research. Special thanks go to my committee chair, Dr. Carroll Hughes, who always managed to find the perfect and humane balance between urging me forward and showing patience while I attended to family matters. I am also most grateful to Dr. Beth Kennard, who is the single person most responsible for my landing at UT Southwestern in the first place. It was a perfect fit for me, and I leave here with the training I came for. I also wish to thank Dr. Graham Emslie for his interest in this project and for generously providing me with access to the data from his CDRR research. Dr. Thomas Carmody provided valuable assistance with statistics and thorough read-throughs, and Dr. Deanna Liss encouraged me to 'think developmentally' about my research. I had a great committee, and I extend my thanks to all of you.

I have come to truly enjoy research, but I am a clinician at heart, and there are many who have helped to shape my clinical identity during my training. First and foremost, I am grateful to my late husband Paul, who was among the finest psychiatrists I have known. His joy, his abundant wisdom, and his humor visit my work with patients. He was pleased that I was doing this, and not a day goes by that I do not long to share the experience with him. Fortunately, a number of

fine clinicians stepped in to fill part of that void. I was fortunate to have two extraordinary supervisors in Gayle Marshall and the late Dr. Diane Myers, who taught me valuable lessons and gave me the confidence to trust my clinical instincts. I also wish to thank Dr. Marie Bannister for the rich clinical dialogue we enjoyed weekly; she managed to supervise and treat me as a colleague without blurring the boundaries. As teachers and clinical mentors, there are few better than Drs. Frank Trimboli, Richard Robinson, and Kathleen Saine. All are equally at home with theory, assessment, and clinical work with patients, and I value (more than they know) all that I learned in their presence. And they made it fun.

My student colleagues provided much appreciated support through the most difficult of times, and I value those friendships. Special thanks go to Lisa Vinuesa-Thoman, Kimberly Doyle, and Jamie Rifkin. Thanks also to my good friends Kelly, Ann, and Sophia Moreno, who 'looked after California' while I was away, and to Marcia Melin, who has been a steady friend through two decades of ups and downs in both our lives. And last but not least, I wish to thank my family, including especially daughters Ann Dee and Alice and my niece Kelly (and all their daughters) for providing the light at the end of the tunnel. Most of all, I am grateful for the loving support of my brother and my late parents—lovely, generous people who instilled in me a love of work and learning. I always hoped my father would live long enough to see me graduate. We almost made it.

August 2008

PATTERN ANALYSIS OF RESPONSE TO ACUTE FLUOXETINE TREATMENT IN THE PREDICTION OF RELAPSE IN CHILDREN AND ADOLESCENTS WITH MAJOR DEPRESSIVE DISORDER

ANN STEVENS AIRY EGGERTSEN, M.S.

The University of Texas Southwestern Medical Center at Dallas, 2008

Supervising Professor: CARROLL W. HUGHES, Ph.D.

Major depressive disorder (MDD) is increasingly recognized as a common and serious affliction among children and adolescents. Antidepressant drug trials aimed at addressing the problem frequently encounter problems in establishing drug efficacy due to the prevalence of placebo effects that are especially prominent in this population, and likewise, placebo responding clouds clinical decisions regarding which patients will benefit from continuation treatment to prevent relapse. Using the method of "pattern analysis," Quitkin and his colleagues (1984, 1987) have shown

that in the acute treatment of depressed adults, true drug benefits are characterized by delayed and persistent improvement, whereas placebo effects tend to occur early and not persist. The present study extends this method to the pediatric population by examining the relationship between Quitkin response patterns during acute fluoxetine treatment and subsequent risk for relapse during randomized placebo-controlled continuation. A total of 168 children age 7 to 18 meeting DSM criteria for MDD first entered 12 weeks of acute treatment on open-label fluoxetine 10–40 mgs with frequent assessment. Using patient response patterns derived from sequential CGI improvement ratings during this period, patients were identified as either true drug or placebo pattern responders in the manner of Quitkin. After 12 weeks, 102 acute responders were randomized to 6 months of continuation treatment on fluoxetine or placebo and monitored for relapse. True drug responders showed an enhanced and very robust fluoxetine-placebo treatment effect (significantly fewer relapses on fluoxetine), whereas placebo pattern responders showed no significant treatment effect. Pattern analysis was also investigated in the larger, multivariate context of predicting risk for depressive relapse in young patients. Cox proportional hazards regression modeling showed continuation treatment and gender to be strong predictors, with the interaction of pattern X treatment falling just short of significance ($p = .07$). Overall results of this study suggest that pattern analysis can be useful in drug studies for pediatric depression and contribute to the prediction of relapse.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF ABBREVIATIONS

ACNP	American College of Neuropsychopharmacology
CBT	Cognitive-behavioral therapy
CDRR	Childhood Depression: Relapse and Remission (study)
CDRS-R	Children's Depression Rating Scale, Revised
CGAS	Children's Global Assessment Scale
CGI	Clinical Global Impressions (scale)
CGI-E	Clinical Global Impressions–Efficacy
CGI-I	Clinical Global Impressions–Improvement
CGI-S	Clinical Global Impressions–Severity
CI	Confidence interval
DSM	Diagnostic and Statistical Manual of Mental Disorders
DSM-III-R	Diagnostic and Statistical Manual of Mental Disorders, Third Edition, Revised
DSM-IV	Diagnostic and Statistical Manual of Mental Disorders, Fourth Edition
DSM-IV-TR	Diagnostic and Statistical Manual of Mental Disorders, Fourth Edition, Text Revision
EPV	Events per variable
FGAS	Family Global Assessment Scale
HAM-D	Hamilton Rating Scale for Depression

K-SADS-PL	Schedule for Affective Disorders and Schizophrenia for School-Age Children–Present and Lifetime Version
MAOI	Monoamine oxidase inhibitor
MDD	Major depressive disorder
NIMH	National Institute of Mental Health
RCT	Randomized controlled trial
SAS®	Statistical Analysis Software
SNRI	Serotonin norepinephrine reuptake inhibitor
SPSS®	Statistical Package for the Social Sciences
SSRI	Selective serotonin reuptake inhibitor
TADS	Treatment for Adolescents With Depression Study
TCA	Tricyclic antidepressant
UTSW	University of Texas Southwestern Medical Center

CHAPTER ONE

INTRODUCTION

Depression was once thought to be a misery reserved for adults. Children and especially adolescents who suffered similarly were generally seen as merely "moody," as if such were a normal phase of development that would resolve naturally over time (Kutcher, 1999; Lefkowitz & Burton, 1978; Weissman & Shaffer, 1998). With the surge of epidemiological studies in the 1980s, pediatric depression came to be recognized as a common and serious public health problem (Weisz, Valeri, McCarty, & Moore, 1999). Major depressive disorder (MDD) is estimated to affect 2.8% of school-age children, and 5.7% of adolescents (Costello, Erkanli, & Angold, 2006), and by 18 years of age, 15−25% of young people have experienced at least one episode of MDD (Birmaher et al., 1996a; Vitiello, Calderoni, & Mazzone, 2006). There is also evidence of a secular increase in prevalence rates, with symptom onset occurring at an earlier age in more recent birth cohorts (Birmaher et al., 1996a; Kutcher, 1999). These numbers represent untold personal costs to patients and their families and place afflicted youth at significant risk for maladaptive functioning as adults (Birmaher et al., 1996a; Dunn & Goodyer, 2006; Kutcher, 1999).

While psychotherapy, especially cognitive-behavioral therapy (CBT), has shown promise in treating depression in young patients (Brent et al., 1997;

Harrington, Whittaker, Shoebridge, & Campbell, 1998; Treatment for Adolescents With Depression Study [TADS] Team, 2004), antidepressant medication is increasingly prescribed as an adjunct or stand-alone treatment (Emslie, Mayes, & Ruberu, 2005). The most commonly prescribed antidepressants are selective serotonin reuptake inhibitors, or SSRIs. In the years between 1989 and 1994, SSRI prescriptions among children and adolescents increased fourfold (Birmaher, Ryan, Williamson, Brent, & Kaufman, 1996b), but for the most part pediatric prescribing predated research. It is only within the last 10 or 12 years that large scale pediatric drug trials have begun to develop a research base to inform antidepressant prescribing practices for the young. In 1997, Emslie and his colleagues (Emslie et al., 1997) published the first major study that convincingly demonstrated the efficacy of the SSRI fluoxetine in the acute treatment of depression in children and adolescents. To date, fluoxetine remains the only antidepressant approved by the Food and Drug Administration (FDA) for the treatment of MDD in young patients (National Institute of Mental Health, 2007; United States Food and Drug Administration, 2007).

Pediatric antidepressant trials face a number of unique challenges, including widespread placebo responding that is often as high as 50% or more (Moreno et al., 2007). When young patients in drug trials improve on active medication, researchers lack the means to differentiate between drug benefits and placebo effects, thus complicating the task of establishing drug efficacy. Further,

even where drug efficacy has been established and significant numbers of young patients achieve remission on antidepressants, the inability to distinguish pharmacologic results from placebo effects leaves open the question of which patients are more likely to benefit from ongoing medication to prevent relapse. Among adult patients, there is evidence that the method of "pattern analysis" can be useful in both these regards (McGrath et al., 2000, 2006; Nierenberg, Quitkin, Kremer, Keller, & Thase, 2004; Quitkin et al., 1987; Quitkin, Rabkin, Ross, & Stewart, 1984).

Pattern Analysis

Studies of antidepressant medications have traditionally relied upon endpoint analysis (at study completion) to compare drug treatment to placebo. The current trend is toward more sophisticated approaches that consider treatment response as a dynamic phenomenon that is best understood when viewed across time. Quitkin and his colleagues at Columbia developed the method of pattern analysis (Quitkin et al., 1984, 1987) to characterize adult response to antidepressant medication during acute treatment. The Columbia group found that among adults, patterns of improvement during initial treatment with antidepressants differentiated between *true drug* response and *nonspecific* or placebo effects (Quitkin et al., 1984, 1987; Rothschild & Quitkin, 1992, Quitkin et al., 1993a). More importantly for the present study, they demonstrated that

these initial response patterns predicted longer term outcomes during ongoing treatment to prevent relapse—that is, 'true drug' pattern responders were found to obtain greater prophylactic benefit from continued medication than others (McGrath et al., 2000; Nierenberg et al., 2004; Stewart et al., 1998). To date, however, no published studies have employed pattern analysis to investigate antidepressant responding in children and adolescents.

The present study proposes to determine if the principles of pattern analysis established by Quitkin and his colleagues in adults can be usefully applied to antidepressant treatment studies in children and adolescents to address the placebo problem in clinical trials, and to investigate the utility of pattern analysis in predicting outcomes during ongoing treatment for relapse prevention.

CHAPTER TWO

LITERATURE REVIEW

Despite occasional articles on the subject in the 1950s and 1960s, interest in early onset depression did not come to the forefront until the 1970s. The National Institute of Mental Health (NIMH) convened its first conference on childhood depression in 1975, bringing to light a simmering controversy over its existence as a valid psychiatric condition (Schulterbrandt & Raskin, 1977). While many called for greater attention to what they considered a serious and long overlooked disorder, others argued that the so called symptoms of childhood depression were little more than transient manifestations of normal development (Lefkowitz & Burton, 1978). As the evolving clinical construct of childhood depression took shape in the 1980s, a wave of epidemiological studies confirmed both its existence and widespread prevalence (Weisz et al., 1999). Treatment studies began to appear with greater regularity, though research on pediatric depression remains relatively sparse compared with adult studies.

Diagnosis of Depression

Early diagnostic formulations of pediatric depression tended to be dynamically oriented around issues of separation and loss (see Schulterbrandt & Raskin, 1977). With successive editions of the Diagnostic and Statistical Manual

of Mental Disorders (DSM), psychiatric diagnosis has become increasingly behavioral, relying more on observable symptom criteria and less on dynamic psychogenesis. The current edition of the Diagnostic and Statistical Manual of Mental Disorders (DSM-IV-TR; American Psychiatric Association, 2000) classifies major depressive disorder (MDD) as a *mood* disorder. Nonetheless, DSM criteria for the diagnosis of MDD reflect a broad constellation of clinical indications that extend well beyond mood. The core criteria are the same for children and adults except for noted differences in how depressive symptoms may manifest in younger patients. According to the DSM, diagnosis of a major depressive episode requires that five of nine symptoms be present daily for two or more weeks and represent a change from previous social or occupational functioning (for children and adolescents, this usually means social withdrawal and/or academic impairment). To reflect the central importance of mood, the 'five or more' symptoms must include either depressed mood or anhedonia; the additional symptoms may include appetite or weight change, sleep disturbance, psychomotor acceleration or retardation, fatigue, feelings of guilt or worthlessness, problems with attention or concentration, and recurrent thoughts of death including suicidal ideation or behaviors. The DSM notes that in children and adolescents, mood may be depressed *or irritable*, and weight loss is broadened to include failure to make expected developmental weight gains. A

number of structured interviews have been developed to assist with diagnosis and to determine course and severity (see Rush, First, & Blacker, 2008).

Treatment Phases

To facilitate research and treatment for major depression, treatment has been conceptualized in three phases (Montgomery, 1994; Vitiello et al., 2006). The *acute* phase of approximately 8 to 12 weeks or more provides primary treatment to address depressive symptoms, aiming to achieve a "response" (clinically significant improvement) or "remission" (elimination of depressive symptoms). Response or remission in the acute phase is often followed by 4 to 9 months of *continuation* treatment to solidify gains and prevent depressive relapse. Those still considered to be at risk may enter a third or *maintenance* treatment phase for one or more years to prevent new (recurrent) episodes of depression. The breakthrough fluoxetine trials by Emslie and his colleagues (Emslie et al., 1997, 2002) addressed the first or acute phase of treatment for pediatric depression. With fluoxetine firmly in place as a first-line medication for the acute treatment of early-onset MDD, attention has turned to the second or *continuation* phase of drug treatment to prevent relapse. Although continuation treatments that combine antidepressant medication and psychotherapy have been shown to be optimally beneficial (TADS Team, 2007), the primary focus of the present study is fluoxetine monotherapy in continuation treatment.

Treatments for Pediatric Depression

Psychotherapy

A wide array of psychotherapeutic approaches to pediatric depression are available and have been described elsewhere (e.g., Weisz et al., 1999), but few have been tested in randomized controlled outcome studies for validation as evidence-based treatments. While there is empirical support for cognitive-behavioral (CBT) and interpersonal (IPT) therapies for young patients (Watanabe, Hunot, Omori, Churchill, & Furukawa, 2007), a recent meta-analytic study found that psychotherapy alone provides only modest benefits for depressed youth, and that the effects are generally not enduring (Weisz, McCarty, & Valeri, 2006). As previously noted, however, results from the multicenter Treatment for Adolescents With Depression Study (TADS Team, 2004, 2007) suggests that a combination of psychotherapy and antidepressants (CBT and fluoxetine) may provide the best treatment.

Pharmacotherapy

Drug trials in children pose ethical issues and require special accommodation, and historically pharmaceutical companies have seen such studies as too costly for the relatively small rate of return (Meadows, 2003). As a consequence, there have been few pediatric drug trials, and until recently most drugs prescribed for children, including psychotropics, were not tested in

children. This resulted in many drugs being prescribed off-label, guided largely by data from adult studies without adequate regard for the differences in pharmacokinetics between children and adults (Meadows, 2003; Vitiello, 2007). The Food and Drug Administration (FDA) Modernization Act of 1997 and the associated Pediatric Research Equity Act of 2003 were enacted to reverse this trend by providing mandates and incentives to encourage pediatric drug research (Meadows, 2003). Prior to this, pediatric antidepressant trials were few.

Soon after tricyclic antidepressants (TCAs) were introduced in the 1960s, clinicans began prescribing them for children and adolescents. Although pediatric drug trials of TCAs were long absent, a dozen studies between 1982 and 1996 indicated that TCAs did not perform significantly better than placebo in treating early-onset depression (Birmaher et al., 1996b; Hazell, O'Connell, Heathcote, Robertson, & Henry, 1995; Kutcher, 1999). Further, TCAs are associated with cardiovascular risks and potentially lethal in overdose.

Another early class of antidepressants, monoamine oxidase inhibitors (MAOIs), were not supported by early evidence for use in children, and were further considered inappropriate for this population due to required dietary restrictions to prevent hypertensive crisis (Kutcher, 1999). To date there have been no placebo-controlled trials of MAOIs in the pediatric population (Wagner, 2005).

The newer SSRI antidepressants came to prominence in the late 1980s. Compared with TCAs and MAOIs, the SSRIs were found to have easier, once daily dosing (due to their longer half-life) and a relatively benign side-effect profile, and physicians began prescribing them for children despite the absence of pediatric trials at the time (Vitiello, 2007). By the 1990s, however, anecdotal observations began linking antidepressants with increased suicidality in young patients (Spotswood, 2004). A modest increased risk for suicidal ideation and behaviors was confirmed by the FDA in a 2004 meta-analysis of 24 randomized controlled trials involving more than 4,500 pediatric patients (results reported in Hammad, Laughren, & Racoosin, 2006). This finding soon led to the FDA mandate for "black box" warnings to monitor suicidality in all children and adolescents being prescribed antidepressants. More recent studies, however, have found a favorable risk-benefit ratio for fluoxetine use in depressed youth (Bridge et al., 2007; Mann et al., 2006), and have determined that the suicide risk is largely mitigated by concomitant CBT psychotherapy (TADS Team, 2007). Further, ecological analysis of regional prescription data (Gibbons, Hur, Bhaumik, & Mann, 2005; Olfson, Shaffer, Marcus, & Greenberg, 2003) has found an overall decline in adolescent suicides as SSRI prescribing has increased. Despite the apparent if slight increase in risk for suicidal ideation among adolescents on SSRIs versus placebo in drug efficacy studies, Olfson et al. (2003) suggest that SSRIs may also mediate the suppression of aggression, thereby reducing lethal

suicidal acts. In any event, the "black box" warning underscores a need to be able to identify and focus antidepressant prescribing on the patients most likely to realize true pharmacologic benefit.

Fluoxetine in Acute Treatment

It is difficult to demonstrate drug efficacy in pediatric antidepressant trials, even where there are substantial drug benefits, because high rates of placebo responding in children tend to obscure results (Kutcher, 1999; Moreno et al., 2007). Despite this difficulty, a breakthrough study by Emslie et al. (1997) demonstrated that the SSRI antidepressant, fluoxetine, is significantly superior to placebo in the acute treatment of depression in children and adolescents. In a double-blind placebo-controlled study, 96 children and adolescents age 7 to 17 were randomized to 8 weeks of treatment with fluoxetine (20 mg) or placebo. Based on the physician-rated Clinical Global Impressions improvement scale (CGI; Guy, 1976), 56% of fluoxetine patients showed significant improvement (*much* or *very much* improved) compared with 33% of placebo patients.

These findings were confirmed in a large multisite study (Emslie et al., 2002) designed to examine the efficacy, safety, and tolerability of fluoxetine in the treatment of pediatric depression. Using a similar research design, 219 depressed young patients age 8 to 18 were randomized to 9 weeks on fluoxetine (20 mg) or placebo. Fluoxetine was found to be superior to placebo across a

variety of outcome measures. This study also recorded any adverse events and found fluoxetine to be safe and well-tolerated in children and adolescents. These two studies by Emslie and his colleagues (1997, 2002) paved the way for fluoxetine to become the first and only FDA-approved drug for the treatment of pediatric depression. Studies examining the efficacy of three other SSRIs—paroxetine, sertraline, and citalopram—have produced weak or conflicting results (Cheung, Emslie, & Mayes, 2006; Hammerness, Vivas, & Geller, 2006; Kane, Fagan, & Wolf, 2007; Mann et al., 2006; Wagner et al., 2003; Wagner et al., 2004). Likewise venlafaxine, a dual-acting serotonin norepinephrine reuptake inhibitor (SNRI), has been shown to be only marginally effective in treating children and adolescents compared with fluoxetine (Emslie, Findling, Yeung, Kunz, & Yunfend, 2007). Despite the established efficacy, relative safety, and widespread use of fluoxetine in pediatric depression, a number of questions remain regarding dosing, treatment duration, relapse prevention, and long-term consequences of treatment.

Fluoxetine in Relapse Prevention

An episode of pediatric depression generally lasts 7 to 9 months (Birmaher et al., 1996a). While the efficacy of fluoxetine for the acute treatment of pediatric MDD is well established (Cheung et al., 2006), less is known about how to treat young patients beyond the acute phase. In order to gain a sense of the rates of

relapse and recurrence in the pediatric population during the post-acute period, Emslie and his colleagues (1998) conducted a naturalistic one-year follow-up study of 87 patients from a previous 8-week placebo-controlled fluoxetine trial. The medication status of patients varied at the one-year mark, but for the group as a whole, 85% of them achieved full recovery from the index episode of depression. Approximately 40% of those who recovered, however, experienced a recurrence at some point during the one-year period, with over half such recurrences occurring within the first 6 months.

Adult research studies have employed a "continuation" design advocated by Quitkin and Rabkin (1981)—also referred to as *discontinuation* studies—to determine whether or not ongoing treatment with antidepressants reduces depressive relapse in adults. Such studies are conducted in two phases: The first is a 2 to 4 month acute treatment phase with antidepressant medication. This is followed by a second and longer continuation phase for which responders are randomized to antidepressant continuation or placebo substitution (i.e., drug discontinuation), generally for an additional 6 to 12 months, to determine if ongoing drug therapy performs better than placebo in preventing relapse.

A continuation study by Montgomery et al. (1988) established the prophylactic efficacy of fluoxetine in adults. A more recent meta-analysis of 31 adult studies (total N = 4,410) by Geddes et al. (2003) found that continuing

drug treatment reduced the risk of depressive relapse by about two thirds across a variety of antidepressants, including MAOIs, TCAs, SSRIs, and others.

In order to examine the potential role for fluoxetine in preventing depressive relapse in children and adolescents, Emslie and his colleagues have conducted two double-blind, placebo-controlled continuation trials with pediatric patients (Emslie et al., 2004, 2008). Patients in both studies were regularly evaluated using the Children's Depression Rating Scale-Revised (CDRS-R; Poznanski & Mokros, 1996) and the CGI, as well as other measures. The first was a small pilot study (Emslie et al., 2004) conducted as part of a larger fluoxetine trial. Researchers analyzed data for 40 young patients age 8 to 18 who had achieved remission (CDRS-R $\leq$ 28) after 19 weeks on fluoxetine (20–60 mg, titrated upward as needed). These remitted patients were randomized to continue on fluoxetine or switch to placebo for 32 additional weeks, and monitored for relapse. Relapse was defined as either CDRS-R > 40 preceded by 2 weeks of clinical deterioration, or relapse in the opinion of the study physician. Results showed that fluoxetine continuation resulted in significantly fewer relapses. Those continuing on fluoxetine relapsed at a rate of 34%, compared with a relapse rate of 60% for those switched to placebo. Fluoxetine continuation also extended the mean time to relapse, which was 181 days on fluoxetine versus 71 days on placebo.

In the subsequent and larger continuation study (Emslie et al., 2008), 102 patients age 7 to 18 who had responded to 12 weeks on fluoxetine (10–40 mg) were randomized to continue on fluoxetine or switch to placebo for a 6-month continuation phase. The degree of clinical improvement required for a patient to advance to the randomized continuation phase was set at CGI ≤ 2 (*much* or *very much improved*), with CDRS-R ≤ 28 or reduced 50% from baseline. Using relapse criteria similar to the previous study (Emslie et al., 2004), investigators again found that fluoxetine continuation resulted in significantly fewer relapses (42%) as compared with placebo (69%). In the absence of fluoxetine, mean time to relapse was 8 weeks, while for patients continuing on fluoxetine, mean time to relapse exceeded 24 weeks.

Taken together, the above two studies suggest that fluoxetine continuation therapy can prevent 4 out of 10 relapses, effectively doubling or tripling mean time to relapse. Although fluoxetine continuation is clearly beneficial in preventing depressive relapse in young patients, there are no established methods for predicting which patients are at risk for relapse, or which ones are most likely to benefit from continuation therapy. Antidepressant research with adults has addressed this question with some success using the "pattern analysis" paradigm.

Pattern Analysis

Patients may improve on antidepressant medication for a variety a reasons. Some improve due to pharmacologic action of the drug itself, while others may improve for nonspecific reasons such as placebo effects or spontaneous remission. Placebo responding while on active medication introduces a confounding factor that complicates the interpretation of results in drug trials. In the early 1980s, researchers set out to find a means to distinguish "true drug" benefits from placebo or other nonspecific effects observed during antidepressant treatment. As a first step, it was necessary to identify response prototypes for patients who improved on active drug versus those who improved on placebo pills. The earliest such studies were conducted by researchers at Columbia University under the direction of Frederic Quitkin.

Pattern Analysis in Adults

In their landmark study, Quitkin and his colleagues (Quitkin et al., 1984) developed and validated the method of pattern analysis to differentiate true drug responding from placebo responding in adults. Fundamentally, their method was to look at depressed adults in placebo-controlled trials to determine if certain aspects of improvement related to *timing* and *stability* (persistence) might differentiate those treated with active drug from those on placebo. They pooled data from three randomized placebo-controlled drug trials involving four

antidepressants: two tricyclics (TCAs), imipramine and desipramine; a monoamine oxidase inhibitor (MAOI), phenelzine; and a tetracyclic, mianserin. After a 10-day placebo washout phase, 185 patients were randomized to a 6-week trial on either drug or placebo. The principle outcome measure was the 7-point Clinical Global Impression (CGI) improvement scale (Guy, 1976), rated weekly for each patient by a study physician. To facilitate pattern detection in a clinically relevant manner, this 7-point scale was dichotomized. Patients who received a weekly improvement rating of GGI = 1 or 2 (*very much improved* or *much improved* compared to baseline) were designated as "improved" for the week, while patients who received a rating of CGI $\geq$ 3 (ranging from *minimally improved* to *very much worse*) were said to be "unimproved." For 111 patients who showed improvement during the 6-week trial, investigators compared sequences ("patterns") of weekly ratings—six for each patient—for those on drug versus those on placebo. As the authors had hypothesized, a pattern of *delayed* and *persistent* improvement occurred more frequently among those on drug than on placebo (41% vs. 15%). That is, for those on active antidepressants, onset of improvement (the initial rating of CGI = 1 or 2) tended to be delayed by 2 weeks or more, and once achieved, to remain stable at CGI $\leq$ 2. This *delayed and persistent* sequence—the most common pattern among those who improved on active medication—became the prototype for "true drug" responding. By contrast, patients who improved on placebo pills were found more likely to show

early onset (within the first 2 weeks) and/or *nonpersistent* improvement. The presence of either attribute accounted for 85% of those on placebo and became the prototype for nonspecific (placebo-type) responding.

The above findings were replicated in a subsequent study (Quitkin et al., 1987) of similar design. Researchers examined response patterns for 85 patients randomized to antidepressants—phenelzine or imipramine—and compared them with 48 patients randomized to placebo. The findings were consistent with the prior study (Quitkin et al., 1984), thus confirming the previously reported response patterns associated with improvement on drug (delayed and persistent) versus improvement on placebo (early and/or nonpersistent).

As a point of clarification, in later pattern analysis studies certain terminology is used interchangeably with the original Quitkin terms for the two acute response pattern types. The *true drug response pattern* is sometimes referred to as a *specific* response (meaning specific to pharmacologic action). Similarly, the *placebo pattern* is often referred to as a *nonspecific* response, which includes placebo effects as well as other nonpharmacologic improvements such as spontaneous remission or natural fluctuation in disease course. Response patterns are also occasionally called *longitudinal* patterns. In all cases, these patterns refer to patient response during the acute phase of treatment.

Pattern Analysis with Fluoxetine

Fieve and others (Fieve, Goodnick, Peselow, Barouche, & Schlegel, 1986) were the first to employ Quitkin pattern analysis in a fluoxetine study. In this double-blind study, 75 mildly to moderately depressed adults were randomly assigned to 6 weeks of treatment on fluoxetine (20, 40, or 60 mg) or placebo. For this study, the CGI-based pattern determination employed a broader criterion of improvement (CGI $\leq$ 3), which—unlike the Quitkin method—also included those only *minimally improved.* Findings were mixed. Among mildly depressed patients who improved, there were no significant pattern differences for those on drug versus placebo. Among moderately depressed patients, however, 27% of those who improved on fluoxetine showed a *persistent* pattern of improvement, whereas none (0%) of those who improved on placebo showed persistence. Despite this latter and seemingly definitive finding, small sample size prevented it from reaching statistical significance. Timing of improvement onset (*early* versus *delayed*), however, did not differentiate between drug and placebo. This may have been due to the more inclusive definition of improvement, likely resulting in too many patients being designated as early responders. In any event, based on this study the usefulness of pattern analysis in drug studies of fluoxetine or other SSRI antidepressants remained unclear.

Four years later, a large multicentered study (Dunlop, Dornseif, Wernicke, & Potvin, 1990) provided clear support for the Quitkin pattern hypothesis in

fluoxetine trials. Among 174 improved patients, a *delayed and persistent* (true drug) response pattern occurred more frequently among those on fluoxetine (31%) than those on placebo, for whom this pattern was markedly rare (8%). Quitkin's prototype patterns again showed the potential to provide researchers with a tool for differentiating between true fluoxetine benefits and placebo effects.

Sensitivity and Specificity

The Columbia group's early studies of pattern analysis (Quitkin et al., 1984, 1987) showed promise for providing a means to sort out the problem of placebo responding in drug trials. Measures of classification accuracy derived from the published data are modest, however. Based on the combined data from 198 responders in these two studies, the correlation between *treatment* (drug vs. placebo) and *response pattern* was moderate ($r = .31$, $p < .001$). For detecting drug responders, pattern analysis showed very good specificity (.81), though the degree of sensitivity (.50) was not as high as researchers would prefer, yielding a number of 'false negatives' (i.e., several responders on active drug were not detected as such). In using pattern analysis to detect placebo responders, the findings were naturally reversed, showing good sensitivity, but low specificity (i.e., a substantial number of 'false positives', indicating patients were on placebo, when in fact they were on active drug). Surprisingly, these limitations have not been included in critical commentary on the Quitkin method of pattern analysis.

Instead, most of its detractors have taken aim at the delayed response component of Quitkin's 'true drug' responding.

Response Onset Controversy

The principle of delayed response to antidepressants is fundamental to the pattern analysis paradigm of Quitkin and his colleagues (Quitkin et al., 1984, 1987). In the Quitkin approach, the "true drug" response to treatment is characterized by *delay* and *persistence*—that is, clinical improvement that appears *after 2 weeks* of treatment, then *persists* through the acute treatment of 6 weeks or more. Some researchers have challenged the *delayed response hypothesis*, asserting that drug effects are evident within days of initiating drug treatment. To a degree, the controversy over the timing of antidepressant onset is a matter of semantics and operational definitions, but a historical review of this issue will shed additional light on the assumptions of pattern analysis, and may suggest possible methodological refinements.

There is published support for both early and delayed antidepressant onset. Initial reports of the first tricyclic antidepressant, imipramine (Kuhn, 1958; Pollack, 1959), indicated that antidepressant action was observable within 1 to 3 days of treatment inception, although occasional patients did not show substantial improvement for 2 to 4 weeks. A randomized controlled trial (RCT) by Small and her colleagues (Small, Milstein, Kellams, & Small, 1981), comparing

imipramine and trazodone with placebo, also reported early antidepressant improvement. In the first and second weeks of treatment, there were greater reductions in mean scores on the Hamilton Depression Rating Scale (HAM-D; Hamilton, 1960) for those on drug than on placebo. By contrast, other RCTs of tricyclic antidepressants (Jones & Ainslie, 1966; Klerman & Cole, 1965) indicated that antidepressant action was delayed, generally appearing after a minimum of 2 weeks of treatment. The landmark pattern analysis studies of Quitkin et al. (1984, 1987) seemed to confirm this finding across a variety of antidepressants, suggesting that early improvements on medication (within 2 weeks) were due to nonspecific placebo effects. Thus, despite prior studies to the contrary (Kuhn, 1958; Pollack, 1959; Small et al., 1981), the Quitkin studies in the mid-1980s established the *delayed antidepressant response hypothesis* as the conventional wisdom among clinicians (Taylor, Freemantle, Geddes, & Bhagwagar, 2006). Within the research community, however, the debate has not been resolved.

Shortly after the Quitkin group's initial pattern analysis studies (1984, 1987), Khan and his colleagues (Khan, Cohen, Dager, Avery, & Dunner, 1989) directly addressed the question of whether early responding reflects drug or placebo effects. They examined data for 129 subjects from three 6-week placebo-controlled trials of imipramine. While the Quitkin group had used the CGI scale

to reflect improvement, Khan et al. employed the HAM-D to gauge relief from depressive symptoms. Early response was defined as a 50% or greater reduction in the HAM-D score with HAM-D $\leq$ 10 at 2 weeks. The findings of Khan et al. supported the delayed response hypothesis. At 2 weeks there was no significant difference between HAM-D subscale scores for drug and placebo treatment groups. A small subset of patients—fewer than one in five—exhibited early response, which was found to be related to the outcome HAM-D score at 6 weeks, but unrelated to treatment (17.5% early responders on drug, 19% on placebo). These findings suggest that early improvement is a nonspecific response unrelated to pharmacologic action. Interestingly, when investigators excluded these early responders from the analysis, the outcome treatment effect (drug vs. placebo) was enhanced, presumably due to a reduction in placebo-type responding that normally introduces troublesome variance into drug trials. Among these delayed responders, the mean HAM-D decline in the first 2 weeks for those on drug ($\Delta_{HAM\text{-}D} = -6.9$) was somewhat greater than for those on placebo ($\Delta_{HAM\text{-}D} = -4.2$), but this statistically significant difference was not large enough to be considered clinically significant. In summary, this study by Khan et al. suggests that early indications of drug action do exist, but are not of sufficient magnitude to meet criteria for early clinical response.

In reviewing two years of research articles, Prien and others (Prien, Carpenter, & Kupfer, 1991) determined that apparent differences in the timing of antidepressant response stem largely from inconsistencies in the terminology and operational criteria associated with clinical change. A companion article by Frank et al. (1991) called for greater precision and consensus in defining key terms such as *improvement*, *response*, *onset of response*, *remission*, *relapse* and *recurrence*. Ten years later, in 2001, the *Journal of Clinical Psychiatry* devoted an entire journal supplement to the topic of early onset antidepressant action. The debate continues today, but one point has become clear: Clinical improvement is a continuous process that researchers often treat as an event that can be pinpointed in time, and the timing of this 'event' hinges on operational considerations—that is, the assessment measures and thresholds of change chosen by researchers to define response.

The Quitkin group employs the Clinical Global Impressions (CGI) improvement rating scale to time the onset of clinical improvement, which results in a relatively high threshold to define response. They explicitly focus on mood and require $CGI \leq 2$ for designating a patient as "improved." $CGI = 2$ is described by them as "at least a 75% reduction in the frequency and intensity of depressed mood" (Quitkin et al., 1993b, p. 563). A majority of antidepressant trials, however, define improvement in terms of percentage reduction in the HAM-D score, variously 20% to 50%, in which case the presence or absence of

early responding depends on where researchers set the bar. In the previously described study by Khan et al., clinical response was defined as 50% or greater reduction in HAM-D score, and thus few early responders were observed.

A Swiss study to examine the relative efficacy of moclobemide and fluoxetine (Stassen & Angst, 1998) was among the first studies to provide separate definitions for antidepressant *response* (50% HAM-D score decline) and *onset of response* (20% score decline that leads to response). Not surprisingly, by using the latter (20%) criterion the rate of early onset response more than doubled as compared with the 50% criterion used by Khan et al. (1989). As in the Khan et al. study, Stassen and Angst also found early response to be predictive of outcome *independent of treatment condition* (drug vs. placebo). Although early HAM-D declines were somewhat greater for those on drug than on placebo, this may be attributable to the sedative-hyponotic effect of many antidepressants which tends to lower HAM-D scores before bringing about improvements in mood. This raises the question of whether these early manifestations of change should be considered pharmacologic side effects or early signs of symptom resolution.

Variable sequence of antidepressant onset. In addressing the debate over timing of antidepressant action, Katz and his colleagues (Katz, Bowden, Berman, & Frazer, 2006; Katz, Koslow, & Frazer, 1997; Katz et al., 2004) have noted that over the years the concept of depression as a clinical entity has evolved from a

unitary illness centering on depressed mood, to a more heterogeneous disorder with multiple symptom components that may include anxiety, psychomotor effects, disturbed sleep and appetite, as well as depressed mood. Further, they cite evidence that various antidepressants target different neurotransmitter systems, inducing different sequences of symptom reduction. Antidepressants may have selective early action that varies from drug to drug. For TCAs and SSRIs, the first week generally brings improvements in insomnia and anxiety, whereas mood and motor symptoms tend to respond more gradually, usually between the second and third weeks. By contrast, the SNRI antidepressant, venlafaxine, appears to initiate rapid improvements in depressed mood and motor functioning with significant effects by day 7, while other symptoms take longer to abate (Derivan & Entsuah, 1995; Rickels, Derivan, Entsuah, Miska, & Rudolph, 1995). Thus, while the neurochemical action of various antidepressants may be almost immediate, their characteristic response profiles—the sequence, timing, and magnitude of specific symptom improvements—may differ.

Onset timing of fluoxetine. A study by Nierenberg et al. (2000) addressed the onset timing of fluoxetine. In this study, 384 depressed outpatient adults received 8 weeks of open-label fluoxetine. Using biweekly HAM-D scores, *onset of response* was defined as a 30% drop in HAM-D score that persisted and led to a *response*, defined in turn as a 50% drop in HAM-D score. Approximately half

the patients (182 of 384) were responders by week 8. Among these responders, over half (55.5%) showed *onset of response* by the end of week 2, with a mean time to *onset* of 3.8 weeks; mean time to *response* itself was 4.9 weeks. These results again suggest that antidepressant effects can be detected early, but a clinically significant response takes longer and is therefore reasonably said to be delayed.

The timing of antidepressant onset continues to be vigorously debated. Posternak and Zimmerman authored a recent article (2005) highly critical of Quitkin pattern analysis, dismissing the delayed antidepressant response theory as clinical lore. To support their position, Posternak and Zimmerman reanalyzed data from 47 placebo-controlled studies collected in a previous meta-analysis by others (Walsh, Seidman, Sysko, & Gould, 2002). At 2 weeks, they found greater mean declines in HAM-D score for those on drug ($\Delta_{HAM\text{-}D} = -7.85$) versus placebo ($\Delta_{HAM\text{-}D} = -5.52$), and asserted that this disproved the delay of antidepressant response. However, as Khan et al. (1989) suggested after obtaining similar results, these early reductions in HAM-D scores for drug and placebo are not sufficiently different to be clinically relevant. Posternak and Zimmerman also noted that for those on drug, 60.2% of the overall HAM-D improvement occurred in the first 2 weeks. Their emphasis on the magnitude of early change ignores the natural course of recovery from depression (seen in HAM-D score trajectories)

that resembles the familiar curve for exponential decay (Priest et al., 1996) commonly seen in medicine and science—steep at first, with declining negative acceleration. Improved sleep alone brings about to a 2 to 6 point reduction of the HAM-D score, but as the pool of remaining symptoms shrinks, with fewer and less tractable symptoms available to remit, the curve 'slows down'.

Quitkin pattern analysis does not assume that there are no detectable mean effects during the first two weeks on antidepressant medication, but rather, that clinically significant changes in mood that result from drug action generally take longer to emerge. Further, the Quitkin method does not rely on mean effects, but on individual patterns that involve both delay *and persistence* based on clinical judgment of patient improvement. The key point here is that even statistically significant early reductions in HAM-D scores during drug treatment are not inconsistent with the principles of pattern analysis. Quitkin pattern analysis clearly has certain shortcomings (e.g., only moderate accuracy), but citing studies that show early detectable response to antidepressant treatment in order to discredit pattern analysis misses the point. Within the pattern analysis paradigm, the focus needs to be on what works best empirically to differentiate drug and placebo effects. Eventually, techniques may be found that outperform the Quitkin method in this regard, but for now it seems clear that this method taps some of the basic mechanisms that underlie patient response to antidepressants.

Measures to define response onset. As previously noted, a key issue in determining when improvement occurs is how terms like *improvement*, *response*, and *response onset* are operationally defined. Curves of symptom reduction over time illustrate that clinical improvement is a gradual and continuous process that is often dichotomized to facilitate data analysis in research—that is, at a given point in time, investigators want to know whether a subject has responded (improved) or not. The answer depends on which measure is used and where the defining threshold is set. The CGI (Guy, 1976) and HAM-D (Poznanski & Mokros, 1996) are commonly used measures, and most adult antidepressant drug trials employ one or both to determine antidepressant response. (Pediatric studies generally use the CDRS-R in lieu of the HAM-D, but the principles are essentially the same.) Though both the CGI and HAM-D (or CDRS-R) are clinician-rated instruments, they vary in approach and emphasis, and the use of one over the other may produce different results regarding antidepressant response.

The CGI improvement scale is a global measure that is adaptable to a wide range of clinical applications to reflect patient status within any specified psychiatric domain (e.g., depression, anxiety, or ADHD). In their use of the CGI, the Quitkin group explicitly emphasized changes in mood over other depressive symptoms and chose a conservative standard of improvement, which they defined as a 75% reduction in depressed mood (Quitkin et al., 1993a). By contrast, the HAM-D and CDRS-R address multiple symptoms of depression, including

somatic symptoms such as insomnia that generally respond rapidly to antidepressant treatment.

As noted above, for the purposes of pattern analysis the measure of improvement onset should be one that effectively distinguishes true drug benefit from nonspecific placebo effects during acute treatment, and in turn contributes to predicting outcome during the continuation phase. No study to date has directly compared instruments of antidepressant response in these two regards. The CGI has two distinct advantages, however, including (a) repeated validation in prior pattern analysis studies, and (b) good ecological validity for treating physicians who are called upon to make rapid global assessments during brief office visits. On the other hand, Katz and his colleagues (1997, 2006) have suggested that existing instruments such as the CGI and HAM-D may both be "too blunt" (or assessed too infrequently) to adequately characterize the sequential unfolding of changes across the multiple symptom domains that respond to antidepressants. Measures that are more sophisticated than the CGI may ultimately prove to be more accurate in identifying patterns that differentiate drug and placebo effects.

Depressive Relapse

Pediatric depression appears to be a chronic and recurrent disorder (Birmaher et al., 1996a). The American Academy of Children and Adolescent Psychiatry (AACAP) reports that 50–75% of children with MDD have a recurrent episode (AACAP, in press, cited in Emslie et al., 2008). Even following successful acute treatment of a depressive episode, approximately 40% of young patients experience a worsening of symptoms or relapse within the first year, more often within the first 6 months (Emslie et al., 1998, 2004). Emslie and his colleagues (2004, 2008) have shown that fluoxetine continuation treatment can reduce the risk of relapse, but little else is known about the mechanism(s) of depressive relapse in the young. Most investigations regarding the nature of depressive relapse have appeared in the adult research literature.

Mechanism of Relapse

As with children, adult studies of depression show that continuation or maintenance medication significantly reduces the subsequent rate of relapse, often to less than half the relapse rate of those switched to placebo (Byrne & Rothschild, 1998). Even with continuation pharmacotherapy, however, studies have reported relapse rates averaging between 7.5% and 25% in adults (Byrne & Rothschild, 1998; Klein, Gittelman, Quitkin, & Rifkin, 1980; Nierenberg et al., 2004; Quitkin et al., 1993b; Stewart et al., 1998; Zimmerman & Thongy, 2007).

Depressive relapse among patients continuing on previously beneficial antidepressants challenges both researchers and treating clinicians. Such relapse may be due to medication noncompliance, increased stress, or illness progression, but the most prominent theory attributes depressive relapse to the development of drug tolerance or "tachyphylaxis" (Cohen & Baldessarini, 1985).

In 1993, Quitkin and his colleagues (Quitkin et al., 1993b) proposed an alternate theory, suggesting that continuation relapse may be due instead to loss of placebo effects. They tested this theory using data from previous placebo-controlled extension studies. A total of 507 depressed adults were first randomized to 6 weeks of acute treatment on imipramine, phenelzine, or placebo. After 6 weeks, 201 patients who were judged to be responders (improvement rating of CGI $\leq$ 2)—166 on drug, 35 on placebo—were continued on the same regimen for an additional 6 weeks. As expected, the rate of relapse for those continuing on placebo (31%) was significantly greater than for those continuing on antidepressants (10%). More importantly, investigators employed the observed rates of response and relapse for drug and placebo treatment groups to estimate how much of the observed relapse could be attributed to loss of placebo effects. Two different assumption models were used. The exclusive model assumed that for those on active drug, pharmacologic action and placebo effects are mutually exclusive, such that patients may be drug or placebo responders, but not both—that is, those who respond to placebo are incapable of a drug response.

The more conservative independent model assumed that improvement may be due to pharmacologic action or placebo effects *or both*, as these mechanisms operate independently and may therefore act simultaneously. The authors reasoned that since the two models "encompass all possible ways" in which placebo effects could interact with drug response to produce improvement, results from the two models would establish the outer limits of the correct answer. For phenelzine, the independent model estimated that 73% of relapse (90% C.I., 36% to 125%) is due to loss of placebo effects, while the exclusive model estimate was over 100%. With 90% probability, then, the 'correct' population estimate for the proportion of phenelzine relapse due to loss of nonspecific or placebo effects is in the range of 36% to 100%. For imipramine, the range was 54% to 100%. Based on these findings, the authors concluded that placebo effects are an important factor in patient relapse during drug continuation.

The above approach was replicated in a recent meta-analytic study looking at newer antidepressants (Zimmerman & Thongy, 2007). The authors used the same *independent* and *exclusive* models to estimate the percentage of relapses attributable to loss of nonspecific effects. Investigators drew data from four continuation trials (total N = 754) involving four different antidepressants including the SSRI paroxetine and three atypical antidepressants—mirtazapine, duloxetine, and nefazadone. Acute treatment durations of 6 or 8 weeks were similar to the prior study by Quitkin et al. (1993b), but the extension phases were

considerably longer, ranging from 20 to 52 weeks (average, 36 weeks). Despite longer continuation periods that extended the opportunity for relapse, the overall relapse rate for patients continuing on antidepressants was relatively low at 7.5%; for those continuing on placebo, it was 24.1%. When the *exclusive* and *independent* models were used to bracket estimates of continuation relapse due to loss of placebo effects, the results were definitive. Estimates for all four drugs approached or exceeded the observed relapse rates. Assuming the *exclusive* and *independent* models (Quitkin et al., 1993b) to be valid, these findings indicate that most if not virtually all of the apparent loss of antidepressant efficacy during continuation treatment is attributable to the loss of nonspecific effects. In other words, the majority of relapsing patients on continuing drug were never 'true responders' to start with.

These two studies (Quitkin et al., 1993b; Zimmerman & Thongy, 2007) strongly suggest that true tachyphylaxis is an insignificant factor in antidepressant relapse during continuation treatment, and that such relapse is more often due to loss of placebo effects (broadly including all nonspecific effects). In turn, it is reasonable to hypothesize that placebo-type responding during acute treatment increases the risk of relapse while on continuing medication. This underscores the potential usefulness of being able to differentiate true drug from placebo pattern responding during initial antidepressant treatment.

Pattern Analysis to Predict Relapse

The above studies suggest that pattern analysis, which differentiates between true drug and placebo response to antidepressants, may be able to contribute to the prediction of relapse. The Quitkin group hypothesized that continued medication would provide greater protection against relapse for true drug responders than for placebo pattern responders, who would be more subject to loss of placebo effects. What follows in this and the following section is discussion of four continuation studies looking at relapse prediction. The first two studies (Stewart et al., 1998; Nierenberg et al., 2004) employ Kaplan-Meier survival analysis to determine whether response patterns predict differential outcomes by treatment. If pattern analysis is valid, these studies would be expected to find a very robust treatment effect (continuing drug vs. placebo substitution) for true drug responders, but little or no such effect for nonspecific, placebo pattern responders. The two succeeding studies (McGrath et al., 2000, 2006) employ regression modeling to examine the predictive validity of pattern analysis in conjunction with other covariates (predictors). These four studies are key to the pediatric study being proposed here.

Stewart and his colleagues (Stewart et al., 1998) were the first to explore the potential for Quitkin pattern analysis to predict relapse. They analyzed data from a previous multisite fluoxetine discontinuation study for which the principal

outcome measure was the 17-item HAM-D. After 12 to 14 weeks on open-label fluoxetine (20 mg), 392 remitted patients (HAM-D $\leq$ 7) entered a year-long continuation phase and were randomized to one of four treatment groups, scheduled to switch from fluoxetine to placebo either (a) immediately, (b) at 3 months (c) at 9 months, or (d) not at all. Relapse was defined as HAM-D $\geq$ 14, or 2 consecutive weeks of meeting MDD criteria per DSM-III-R (American Psychiatric Association, 1987). Survival analysis at 3 months into the continuation phase showed that among *true drug pattern* responders, relapse on placebo substitution occurred at more than twice the rate (54%) of those still on fluoxetine (23%). By contrast, for *placebo pattern* responders, rates of continuation relapse did not differ significantly for fluoxetine versus placebo. (Treatment effects were diminished at the 9-month and one-year assessments.) These results demonstrated that pattern analysis can make a significant contribution to predicting relapse in adults, and suggested that the benefits of fluoxetine continuation may be limited to patients with a true drug pattern of initial (acute) response. The findings of this study are consistent with the previously described studies (Quitkin et al., 1993b; Zimmerman & Thongy, 2007) demonstrating that relapse during continuation treatment is largely due to loss of placebo effects rather than tachyphylaxis.

A subsequent placebo-controlled discontinuation study of mirtazapine (Nierenberg et al., 2004) sought to replicate the above fluoxetine findings. This

multisite study again examined the relationship between acute patterns of response and continuation relapse. During acute treatment, 410 depressed adults received 8 to 12 weeks of open-label mirtazapine treatment (15–45 mg, titrated upward as needed). At the conclusion of the acute phase, 156 remitted patients (HAM-D $\leq$ 7 and CGI-I $\leq$ 2) entered a 40-week double-blind continuation phase to be randomized to mirtazapine continuation or placebo substitution. Relapse was subjectively defined as clinical deterioration requiring a change in treatment in the judgment of study investigators. Acute longitudinal response patterns were again found to contribute to predicting continuation outcomes. As hypothesized, those with a "true drug initial" pattern of response to mirtazapine (delayed and persistent improvement) clearly benefited from ongoing drug treatment, experiencing a significantly lower rate of relapse on drug (25%) than on placebo (56%). Also as hypothesized, patients with a "placebo initial" pattern of response to mirtazapine (early and/or nonpersistent improvement) did not show significantly different rates of relapse for mirtazapine versus placebo continuation treatment, although there was an unexpected trend suggesting that ongoing drug treatment provided these patients with *some* protection against depressive relapse (14% relapse on drug continuation vs. 30% on placebo, $\chi^2 = 2.1$, $df = 1$, $p = .14$). Still, the main conclusion from these two discontinuation studies (Nierenberg et al., 2004; Stewart et al., 1998) is that adult patients who exhibit true drug response patterns are at greater risk of relapse if antidepressant medication is prematurely

withdrawn, and are the patients who stand to benefit most from continuing pharmacotherapy.

Although both of these studies support the predictive validity of pattern analysis with regard to continuation relapse, they vary with regard to the benefits of drug continuation for placebo pattern responders. As hypothesized, in the earlier fluoxetine study by Stewart et al. (1998), placebo-type responders show no significant benefit for fluoxetine continuation over placebo substitution. In the mirtazapine study by Nierenberg et al. (2004), however, placebo pattern responders showed a modest if nonsignificant trend indicating that drug continuation may afford a modicum of protection against relapse. Nierenberg et al. suggest that the apparent discrepancy between studies may stem from differences in study populations. Patients in the earlier fluoxetine study (Stewart et al., 1998) tended to be more severely or chronically ill and therefore more subject to relapse, which may have obscured modest differences between drug and placebo treatment outcomes for placebo pattern responders. In any event, it is possible that the predictive power of pattern analysis partly depends on certain clinical attributes of the patients as, for example, the severity or chronicity of the depressive illness. Nierenberg et al. (2004) also point to the inexact and heterogeneous nature of response patterns, such that even apparent placebo-type responders may show some pharmacologic benefit from drug continuation

(providing support for the *independent* model in Quitkin et al., 1993b), though the effect size is clearly greater for true drug pattern responders.

The above studies (Nierenberg et al., 2004; Stewart et al., 1998) support the predictive validity of pattern analysis, and suggest that antidepressant continuation therapy is especially useful in those who exhibit a *true drug* initial response pattern. These studies also lend further support to the hypothesis that a significant portion of the relapse observed on continued medication is due to loss of nonspecific (placebo) effects as opposed to loss of drug effects (tachyphylaxis).

Relapse Modeling

Researchers have attempted to determine which patient-specific factors predispose pediatric patients to depressive relapse (e.g., severity of the index episode, comorbid disorders, family dysfunction, or demographic factors). Studies to date have produced inconsistent results (Emslie, Mayes, & Ruberu, 2005), and except for the protective effects of fluoxetine continuation recently demonstrated by Emslie et al. (2004, 2008), other predictors of relapse among children and adolescents remain elusive. Adult studies have had modest success in this area.

In order to look beyond pattern and treatment to examine other potential predictors of risk for relapse, the Quitkin group (McGrath et al., 2000) applied multivariate regression modeling to data from a previous fluoxetine discontinuation study (Stewart et al., 1998, described above) in which acute

treatment consisting of 12 to 14 weeks on open-label fluoxetine (20 mg) was followed by randomized discontinuation (placebo substitution) occurring either immediately, at 3 or 9 months, or not at all. Consistent with previously published survival analysis of this data, regression modeling confirmed that response pattern and continuation treatment interacted significantly to predict time to relapse—that is, drug-placebo differences in continuation outcomes were seen only in those with specific (true drug) patterns. The analysis also found two additional variables of interest. The presence of positive neurovegetative symptoms (insomnia, loss of appetite) showed a very robust advantage for fluoxetine continuation over placebo, but only in those with specific (true drug) longitudinal response patterns. For patients with reverse neurovegetative symptoms (hypersomnia, weight gain, lethargy), fluoxetine continuation provided no significant protection against relapse. Further, chronic depression ("chronicity")—whether signaled by early onset of depression (before age 20) or episode duration (more than 2 years)—increased the risk of relapse during the continuation phase by 40%, although there were conflicting statements regarding whether treatment effect was differential by chronicity. Other variables including age, gender, and history of hypomania were not found to be predictive of differential survival (nonrelapse) between fluoxetine and placebo during the continuation phase. This study was important in showing that pattern analysis not only provides a useful research tool, but may ultimately fit into a clinical decision model to determine

the need for continuation treatment in remitted patients. Based on their findings, McGrath et al. (2000) demonstrated this potential clinical utility in this excerpt:

> For example, patients with an early onset of depression, typical vegetative symptoms, and a specific [true drug] longitudinal pattern could be advised strongly that, for them, the advantage of fluoxetine continuation over placebo is large (60%, 95% CI = 42% to 79%) and continuation / maintenance warrants great efforts to control side effects. Other [remitted] patients without any of these 3 features have little or no discernible advantage from continued fluoxetine treatment, at least at a fixed daily dose of 20 mg. (p. 523)

A subsequent fluoxetine study by the Quitkin group (McGrath et al., 2006) was the first prospective study of depressive relapse designed to test the predictive validity of response patterns and other potential predictors of continuation relapse. In addition to pattern analysis and continuation treatment, researchers examined multiple covariates including age, gender, neurovegetative symptoms (positive or reverse), chronicity, severity, depressive subtype, and magnitude of acute-phase response. During the acute treatment phase, 570 patients received 12-weeks of treatment on open-label fluoxetine, titrated upward from 10 mg to 40–60 mg as tolerated. Following acute treatment, 292 responsive patients were randomized to continue on the same dose of fluoxetine or to switch to placebo for 52 additional weeks (or until relapse). Baseline and post-acute phase depression severity were assessed using a modified HAM-D to account for reverse neurovegetative symptoms when present. Treatment response was defined in terms of HAM-D difference scores ($\Delta_{HAM\text{-}D}$) and CGI improvement ratings. Chronicity was

measured on a 6-point scale, varying from 1 = *single depressive episode*, to 6 = *chronic, persistent depression*. Results showed that fluoxetine continuation significantly reduced the rate of relapse compared with placebo, but contrary to prior studies (McGrath et al., 2000; Nierenberg et al., 2004; Stewart et al., 1998) pattern analysis did not interact with treatment to predict relapse. The authors were unable to explain this unexpected departure from earlier findings. Acute treatment durations were equivalent in the two studies. In the more recent study (McGrath et al., 2006), upward dose titration (perhaps affecting response patterns) and a low threshold of response for advancement to the continuation phase (resulting in the inclusion of weak responders to fluoxetine) may have reduced statistical power. The authors suggest that only independent replication can resolve the issue of the predictive validity of pattern analysis. Other factors that were found to predict greater risk of relapse included female gender, reverse neurovegetative symptoms, pretreatment severity, minimal response at time of randomization, and chronic history of depression. Chronic depression affected nearly 60% of patients in this study and, as in the prior study (McGrath et al., 2000), chronicity was found to be highly predictive of relapse independent of continuation treatment. (Indeed, the prevalence of chronic depression in this study cohort may have overshadowed some elements of the prediction model, including differential treatment effects by pattern.) Based on the strength of chronicity as an independent predictor of relapse, the authors recommended that

future continuation or maintenance studies be stratified on this factor to avoid confounding effects.

Nature of Placebo Effects

"Depression is an inherently placebo-responsive condition."
(Khan & Brown, 2001, p. 123)

The development of pattern analysis was largely driven by the troublesome problem of placebo effects in evaluating drug efficacy, whether in drug trials or clinical practice. The term "placebo effect" commonly refers to symptom relief on sugar pills or other inactive agents that results from patients' expectations. In antidepressant drug trials, the term more broadly includes any improvement that is nonpharmacologic in nature, regardless whether treatment is active drug or placebo. Although optimism and expectation clearly contribute to placebo effects, the term *placebo* comes from the Latin, "I shall please" (Moreno et al., 2007). Thus, it is not surprising that children, who naturally seek attention and approval from adults, are especially prone to placebo responding. In pediatric antidepressant trials, the average rate of response to placebo pills approaches 50%, as compared with about 30% in adults (Moreno et al., 2007). This is problematic for pediatric drug trials, where high rates of placebo responding—to both active drug and placebo pills—decrease treatment effect size by reducing the signal-to-noise ratio, rendering the detection of true pharmacologic benefit more

difficult (Khan & Brown, 2001). The cost of the 'placebo problem' may be evident in the profusion of failed pediatric trials. A recently published article from the American College of Neuropsychopharmacology (ACNP, Mann et al., 2006) reported on placebo response rates for 11 pediatric antidepressant trials, which averaged around 50%. Notably, placebo response rates for the three successful fluoxetine trials included in the report averaged closer to 40%. While having an effective drug is key to establishing drug efficacy, methods to reduce or at least mitigate placebo responding in pediatric trials are essential. As drug trials in children and adolescents are relatively recent, it is again necessary to turn to the adult literature to review what is known about the nature of placebo effects.

There appears to be a tacit assumption that pill taking is instrumental in placebo responding. This notion was examined in a study of pretrial placebo washout responders (Rabkin et al., 1990). Fifty mildly depressed adult patients who showed marked improvement during a 10-day placebo washout were randomly assigned to 6 weeks of continued or discontinued placebo medication (i.e., pills or no pills). Relapse rates for the two groups were identical (52%), suggesting that pill taking itself may not be a critical element in maintaining a placebo response. Beyond pill taking, placebo responding in drug trials may also be associated with nonspecific effects that derive from attention and social support inherent in the caregiving environment. Further, some clinical

improvement may simply result from regression to the mean (Dworkin, Katz, & Gitlin, 2005)—that is, the course of depression naturally fluctuates, and patients who tend to seek treatment at the nadir of their depression also tend to get better.

In a pair of companion studies, the Quitkin group employed pattern analysis to further explore the nature of placebo responding (Quitkin, McGrath et al., 1991; Quitkin, Rabkin et al., 1991). While earlier pattern analysis studies had focused on just two aspects of improvement—*time of onset* (early versus delayed) and *persistence*—these two 1991 studies added a third dimension, *rate of onset*. Rate of improvement onset was characterized as "abrupt" or "gradual" based on weekly CGI improvement ratings and whether or not the patient 'abruptly' skipped a CGI ratings category just prior to improvement—for example, going from CGI = 4 directly to CGI = 1 or 2. Investigators employed data accumulated in a number of placebo-controlled antidepressant trials with moderately to severely depressed adults involving a variety of antidepressants. The first of these two studies (Quitkin, Rabkin et al., 1991) examined data for 144 patients randomized to treatment with placebo, while the companion study (Quitkin, McGrath et al., 1991) focused on 263 patients randomized to acute treatment with active drug. The findings of the two studies were complex, but fully consistent with the investigators' developing understanding of response pattern dynamics. Based on their results, the authors concluded that acute response to placebo is *heterogeneous*, representing at least two different mechanisms of response:

(a) *Abrupt* improvement on placebo, which tends to be early but fleeting (nonpersistent), seems to indicate placebo effects resulting from an initial surge of optimism. (b) On the other hand, *gradual* improvement on placebo, which tends to be more stable (persistent), is consistent with the natural course of spontaneous remission. The authors found that among those on active drug, just as for those on placebo, abrupt change that occurred early tended not to persist, again suggesting a placebo effect. While late occurring abrupt improvement (after 2 weeks) in patients on active medication was rare, such improvement persisted 75% of the time, perhaps suggesting the abrupt event was associated with the drug 'taking hold'. By contrast, all abrupt improvement that occurred late in placebo patients was transient. In summary, the implications of abrupt improvement depend on treatment and timing, but whether on drug or placebo, the observation of abrupt improvement early in the acute phase strongly suggests placebo effects.

Placebo-type responding is an ongoing problem in antidepressant drug trials. With good research design and sophisticated data analysis, it has not been inordinately difficult to establish antidepressant efficacy in adults, but higher rates of placebo-type responding in pediatric trials present more of a challenge. It is worth noting that the placebo problem in antidepressant trials appears to be growing (Walsh et al., 2002). For a discussion of possible reasons (including changes in recruitment procedures and study populations) as well as suggestions for methods to address the problem, see Dworkin et al. (2005). Regardless of the

reasons for the increase in placebo-type responding, pattern analysis takes on increasing importance as a potential tool for detecting placebo effects.

The literature reviewed here summarizes prior work highlighting two current problems that confront antidepressant research and treatment in pediatric populations. The first is the high rate of placebo responding that confounds drug efficacy trials. The second problem is the lack of reliable means to determine patient risk for depressive relapse, leaving treating clinicians with few guidelines regarding continuation treatment in remitted patients, where the goal is to focus drug treatment in those youth who will obtain true pharmacologic benefit. The present study will address these issues by applying pattern analysis methodology to children and adolescents who participated in a large NIMH study examining the role of fluoxetine in the prevention of depressive relapse. Specifically, this study proposes to test the predictive validity of pattern analysis with regard to relapse in pediatric depression.

CHAPTER THREE

STUDY AIMS AND HYPOTHESES

Aims of Study

Large-scale antidepressant drug research in children did not begin in earnest until the late 1990s, and by comparison with research in adults, relatively few studies have been published. Studies by Emslie and his colleagues (Emslie et al., 1997, 2002) established the efficacy of fluoxetine in treating pediatric depression, but much remains to be learned about the nature of the antidepressant response in young patients. Placebo response in patients on active antidepressants is a troubling source of variance in most drug trials, and can be especially problematic in pediatric studies where placebo effects are more prominent (Moreno et al., 2007). It would be useful, therefore, to be able to distinguish between placebo and true drug responses in children and adolescents. Further, despite the widespread use of fluoxetine to treat pediatric depression, there are few guidelines to suggest which young patients are most likely to benefit from continuation therapy to prevent relapse. To address these issues, the present study extends the Quitkin method of pattern analysis to children and adolescents in a large NIMH-funded trial investigating the efficacy of fluoxetine in the prevention of relapse in pediatric depression (Emslie et al., 2008). The overarching goal here is to investigate the Quitkin method of pattern analysis as a means to differentiate

true drug from placebo effects in the pediatric population. The specific aims of the present study are as follows:

1. To determine whether pattern analysis contributes to the prediction of relapse among young patients who respond to acute treatment with fluoxetine.
2. To investigate other clinical and demographic variables that may predict relapse, either independently or in interaction with pattern analysis and continuation treatment.

Successful distinction between true drug and placebo pattern responders in pediatric patients on fluoxetine has the potential to clarify the results of clinical trials and to inform clinical decisions regarding continuation treatment for the prevention of relapse.

Primary Hypotheses

This study examines whether or not pattern analysis of patient response to an initial, 12-week (acute) phase of fluoxetine treatment in children and adolescents predicts outcome (relapse vs. continued remission) for a 6-month continuation phase in which patients are randomized in double-blind fashion to fluoxetine versus placebo. Toward this end, four primary hypotheses (after Nierenberg et al., 2004; Stewart et al., 1998) were studied:

Hypothesis 1

Patients with a *true drug* pattern of response to acute fluoxetine treatment (*delayed* and *persistent* improvement) will be more likely to relapse on placebo than fluoxetine during the randomized continuation treatment.

Hypothesis 2

For patients with a *nonspecific/placebo* pattern of response during acute fluoxetine treatment (*early* or *nonpersistent* improvement), relapse during the continuation phase will not depend on continuation treatment (drug vs. placebo).

Hypothesis 3

For patients switched to placebo during randomized continuation treatment, those who previously exhibited a *true drug* pattern of response to fluoxetine will relapse at a higher rate than those who previously exhibited a *nonspecific/placebo* pattern of response to fluoxetine.

Hypothesis 4

For patients randomized to fluoxetine during continuation treatment, those who previously exhibited a *true drug* pattern of response will have a lower rate of relapse than those who previously exhibited a *nonspecific/placebo* pattern of response to fluoxetine. Stewart et al. (1998) found support for this hypothesis, while Nierenberg et al. (2004) did not. A positive finding would support 'loss of

placebo effects' as a more prominent explanation of relapse during continuation treatment (as an alternative to tachyphylaxis).

As noted by Stewart et al. (1998), any demonstration that short-term response patterns alter subsequent prognosis serves to validate pattern analysis.

Secondary Hypotheses

Researchers have previously attempted to determine which patient-specific factors predispose young patients to relapse, but with inconsistent results (Emslie, Mayes, & Ruberu, 2005). This study employs regression analysis to model relapse, employing multiple covariates as predictors, starting with but not limited to pattern and continuation treatment.

Hypothesis 5

In conjunction with acute response patterns and continuation treatment, other demographic and/or clinical factors, such as age, gender, depression chronicity and severity, neurovegetative status, and degree of acute response will contribute to predicting time to relapse during continuation treatment. Based on prior findings in adults (McGrath et al., 2000, 2006; Stewart et al., 1998), the following specific findings are anticipated:

Hypothesis 5a. Acute response pattern will interact with continuation treatment to enhance the prediction of time to relapse. A positive finding here will support the primary hypotheses (1–4).

Hypothesis 5b. Neurovegetative status will interact with continuation treatment to enhance relapse prediction. Specifically, positive neurovegetative symptoms (insomnia, loss of appetite) will increase the effect of fluoxetine in reducing risk for relapse. For these with reverse symptoms (hypersomnia, weight gain), fluoxetine will confer no more benefit than placebo in preventing relapse.

Hypothesis 5c. Chronicity of illness will predict outcome (relapse vs. remission) independent of treatment. As a corollary, pattern analysis will be more effective in predicting differential response to fluoxetine versus placebo in nonchronically depressed patients.

Exploratory Hypotheses

Hypothesis 6

Acute response patterns will be associated with differential dropout rates on fluoxetine continuation versus placebo. Specifically, true drug responders will drop out sooner and more frequently when randomized to placebo substitution (vs. fluoxetine). Conversely, placebo-type responders will not show differential dropout rates for fluoxetine and placebo. This has not been explored in the adult literature, and is proposed as an a priori hypothesis that is consistent with what is

known about pattern analysis. A positive finding will help to inform the interpretation of results from Hypotheses 1–5.

Hypothesis 7

True drug and placebo pattern responders will differ on some demographic and clinical attributes. In the absence of prior research, this is an exploratory hypothesis to determine differential characteristics, if any, between true drug vs. placebo pattern responders.

Hypothesis 8

Abrupt improvement (e.g., a jump in global improvement rating from CGI-I = 4 to CGI-I = 1 or 2 in consecutive weeks/assessments) during the acute phase will tend to occur within the first 2 weeks, and if so, not to persist (see Quitkin, McGrath et al., 1991; Quitkin, Rabkin et al., 1991). As such, early abrupt improvement would provide an added marker for placebo-type responding.

Hypothesis 9

A six-digit acute response pattern based on the first 8 weeks of acute treatment (weeks 1, 2, 3, 4, 6, 8) will relate to outcome as effectively as the eight-digit pattern reflecting the full 12 weeks of treatment (weeks 1, 2, 3, 4, 6, 8, 10, 12). This hypothesis acknowledges the gradual evolution of pattern analysis over 20 years of research. Pattern analysis was initially validated in adults by using

weekly CGI improvement ratings during a 6-week acute treatment phase. As acute drug trials have expanded to 8, 10, 12, or even 14 weeks (McGrath et al., 2000; McGrath et al., 2006; Nierenberg et al., 2004; Stewart et al., 1998)., the Quitkin pattern method has been modified to accommodate the longer treatment durations, often accompanied by biweekly (vs. weekly) ratings during the latter weeks of acute treatment. Despite this, there are no known studies validating these modifications to the original Quitkin approach.

This study is intended to provide a preliminary examination of pattern analysis as a potentially useful methodology in antidepressant research on pediatric depression. The developing research literature suggests that the Quitkin pattern analysis approach may oversimplify a response that almost surely involves a complex array of symptom changes that unfold sequentially and continuously during treatment. The overarching goal of this study is not so much to validate the Quitkin method per se. Rather, the aim here is to test the validity of the underlying premise of pattern analysis in pediatric depression—that young patients experiencing true drug benefits versus placebo effects respond to antidepressant treatment in detectably different ways. These differences (patterns of acute response) have the potential to usefully inform research and continuation treatment in pediatric depression. A showing of even modest validity in children and adolescents would provide support for future research aimed at refining these techniques to increase the predictive validity of pattern analysis.

CHAPTER FOUR

METHODS

The present study is based on data collected as part of a larger pediatric depression study, Childhood Depression: Remission and Relapse (CDRR), conducted at the University of Texas Southwestern Medical Center at Dallas (Graham Emslie, principal investigator, 2000). The CDRR study was funded by NIMH to examine the efficacy of fluoxetine, an SSRI antidepressant medication, in preventing depressive relapse in children and adolescents during continuation treatment (Emslie et al., 2008; Mayes et al., 2007). This randomized controlled discontinuation trial took place between August 2000 and July 2006. Qualifying patients received 12 weeks of acute treatment on open-label fluoxetine, after which responders were randomly assigned to 6 months of fluoxetine continuation or discontinuation (placebo) and monitored for relapse.

Participants

Subjects for the CDRR study were children and adolescents with depression who were recruited from clinical referrals to the Child and Adolescent Psychiatry Outpatient Clinic at the Children's Medical Center of Dallas.

Inclusion Criteria

Study participants included outpatients between the ages of 7 and 18 years with a primary diagnosis of nonpsychotic major depressive disorder (MDD) meeting DSM-IV criteria for a minimum of 4 weeks. Diagnosis was established using the Schedule for Affective Disorders and Schizophrenia for School-Age Children–Present and Lifetime Version (K-SADS-PL; Kaufman et al., 1997). Depression severity was required to be at least moderate, as determined by the Children's Depression Rating Scale, Revised (CDRS-R $\geq$ 40; Poznanski & Mokros, 1996) , and the Clinical Global Impressions–Severity scale (CGI-S $\geq$ 4; Guy, 1976). Participants were required to be of normal intelligence (IQ $\geq$ 80), attending school, and deemed medically healthy. For prospective patients of questionable intellectual capacity, provision was made for administration of the Wechsler Intelligence Scale for Children–Third Edition (WISC-III; Wechsler, 1991).

Exclusion Criteria

Patients were excluded from the study if they had a history of bipolar disorder, schizophrenia or other psychotic disorder (including psychotic depression), anorexia or bulimia nervosa, or alcohol or substance abuse in the previous 6 months. Certain other comorbid conditions were allowed, including anxiety disorder, attention deficit/hyperactivity disorder (ADHD), or conduct

disorder. Pregnant and lactating females or sexually active females not employing effective methods of contraception were excluded. Patients with chronic medical conditions requiring regular use of medications with psychotropic side effects (e.g., anticonvulsants, steroids) were excluded, although ADHD patients who were stable on psychostimulant medication or Strattera® (atomoxetine) were not excluded. Also excluded were patients with first order relatives with bipolar I disorder, and patients with past histories of serious suicidal ideation or behavior. Subjects were excluded it they had previously failed an adequate trial on fluoxetine or had shown intolerance to fluoxetine.

Informed Consent

The CDRR study was approved by the Institutional Review Board (IRB) of the University of Texas Southwestern Medical Center. All participants and their parents or guardians were apprised of the purpose and procedures of the study, including potential benefits and risks, and their rights as patients. Prior to enrollment, parents signed written informed consent, and participants provided their assent.

Procedures

Screening

Patients referred to the CDRR study were screened by trained interviewers (research assistants) for possible inclusion, and potential subjects were scheduled for a diagnostic interview. After informed consent was obtained, the parent(s) and child were separately interviewed using the K-SADS-PL to establish diagnosis and course of illness. Child functioning was assessed using the Children's Global Assessment Scale (CGAS; Shaffer et al., 1983), and family functioning was assessed with the Family Global Assessment Scales (FGAS; Mrazek & Masterson, 1985). Family psychiatric history was documented using the Family History Research Diagnostic Criteria (FHRDC; Andreasen, Rice, Endicott, Reich, & Coryell, 1986). Patients who appeared to meet study criteria underwent further evaluation one week later by a psychiatrist or licensed psychologist. At that time, course of illness and depressive severity were confirmed using the K-SADS-PL, CDRS-R, and CGI-S scale. Patients who continued to meet all study criteria were enrolled in the acute treatment phase of the study. Participants commenced acute treatment within 5 to 10 days following baseline evaluation.

Acute Treatment

During the 12-week acute phase, patients were treated with open-label fluoxetine. They received a daily dose of 10 mg of fluoxetine for the first week, after which the dose was increased to 20 mg. After 6 weeks, dosing could increase to 30 or 40 mg if a patient showed minimal or no improvement (CGI $\geq$ 3). For patients with bothersome side effects, dosing could be reduced to 10 mg.

Participants came in for regularly scheduled visits with a child psychiatrist who assessed patient status, adjusted fluoxetine dosage as necessary, and completed all rating scales based on child and parent interviews. Ratings included the CDRS-R, CGI, CGAS, and FGAS. Participants were seen weekly for the first month (weeks 1, 2, 3, 4), then every other week thereafter (weeks 6, 8, 10, 12). No formal psychotherapy was permitted apart from supportive clinical management. Patients were removed from the study if pill counts reflected noncompliance, that is, using less than 70% of pills for two consecutive visits or a total of three visits. At the end of 12 weeks, nonresponders (CGI $\geq$ 3, indicating little or no improvement, or worsening of symptoms) were discontinued from the study and provided with treatment recommendations and referrals as needed.

Continuation Phase

After 12 weeks of acute treatment, participants were eligible to enter the 6-month continuation (randomization) phase if they were in remission (CGI-I $\leq$ 2

and CDRS-R $\leq$ 28) or showed adequate clinical response (CGI-I $\leq$ 2 and CDRS-R score decline of 50% or more). Eligible patients who consented to continue were randomized in double-blind fashion to either continue fluoxetine at the acute treatment dose, or switch to placebo. Computer randomization stratified patients by gender, age (12 years and younger vs. 13 and older), and acute treatment response (remission vs. adequate clinical response). For the first month of continuation treatment, patients visited a study psychiatrist for evaluation every other week (weeks 14, 16), then every 4 weeks thereafter (weeks 20, 24, 28, 32, 36). Rating instruments administered at these visits again included the CDRS-R, CGI, CGAS, and FGAS. As during the acute phase, medication noncompliant patients were removed from the study. Further, patients who relapsed or whose welfare might be jeopardized by continued participation were removed from the study and referred for appropriate treatment as needed.

Definition of relapse. Based largely on the CDRS-R, and in keeping with definitions established for the CDRR study, relapse was defined as either (a) a one-time score of CDRS-R $\geq$ 40 with worsening depressive symptoms for 2 weeks or more, or (b) clinical deterioration suggesting relapse would be likely without altering treatment, even if CDRS-R < 40. Either condition was considered "relapse" for the purpose of primary data analyses in this study. Some secondary analyses were conducted using only the stricter component of the definition above (CDRS-R $\geq$ 40), referred to hereafter as "full relapse."

Measures

While multiple assessment measures were administered in the course of the CDRR study, only the measures relevant to the current study are described here.

K-SADS-PL

The Schedule for Affective Disorders and Schizophrenia for School-Age Children–Present and Lifetime Version (Kaufman et al., 1997; Kaufman, Birmaher, Brent, Rao, & Ryan, 1996) is a comprehensive semi-structured diagnostic interview designed to assess current and past history of psychopathology. The K-SADS-PL yields diagnostic information based on separate interviews with child and parent conducted by a psychiatrist or other qualified clinician. The K-SADS-PL includes a screening interview that addresses 82 symptoms across 20 diagnostic areas. Depending on results of the screen, the interviewer may be directed to one or more of five supplements to address differential diagnosis within key diagnostic domains that include affective disorders; psychotic disorders; anxiety disorders; behavior disorders; and substance abuse, eating, and tic disorders. The K-SADS-PL provides probes (questions) and objective criteria to aid the interviewer in eliciting sufficient information to score individual items relating to current and past episodes. The majority of items are scored using a rating scale of 0 to 3, where 0 indicates no information is available, 1 indicates symptom absence, 2 suggests subthreshold

symptoms, and 3 represents clear symptom presence. Based on data synthesized from parent and child interviews, the K-SADS-PL is able to determine patient status regarding 32 DSM-IV child psychiatric diagnoses which can each be designated as definitive, probable, or not present. It also provides both global and diagnosis-specific levels of impairment. In psychometric studies by Kaufman and her colleagues (Kaufman et al., 1997), interrater agreement ranged from 93% to 100% for both present and lifetime diagnoses, with kappa coefficients ranging from .63 to 1.00 for various current and past diagnoses. Diagnostic stability for intervals of 1 to 5 weeks were good to excellent across most current and lifetime diagnoses, with test-retest reliability coefficients of .86 to 1.00 for mood disorders. The K-SADS-PL interview is estimated to require 35 to 75 minutes depending on the degree of psychopathology. In the present study, the K-SADS-PL was the principal instrument employed to establish DSM-IV Axis I diagnoses prior to enrollment.

CDRS-R

The Children's Depression Rating Scale–Revised (Poznanski et al., 1984; Poznanski & Mokros, 1996) is a clinician-rated instrument designed to assess the severity of depressive symptoms in children. It is sensitive to medication effects and has become the most common measure for tracking clinical outcomes in studies of child and adolescent depression (Pavuluri & Birmaher, 2004). Modeled

after the Hamilton Depression Rating Scale (HAM-D) for adults, it is a semi-structured interview that assesses 17 symptoms associated with DSM-IV criteria for MDD. The CDRS-R is administered separately to the patient and parent(s) or other supplementary informant such as a teacher. The 17 interview items (symptoms) are accompanied by descriptive rating anchors to facilitate accurate scoring. Fourteen of the CDRS-R items are based on verbal responses related to impaired schoolwork, difficulty having fun, social withdrawal, sleep disturbance, appetite disturbance, fatigue, physical complaints, irritability, guilt, low self-esteem, depressed feelings, morbid ideation, suicidal ideation, and weeping. The final three items reflect interviewer observations regarding language tempo, hypoactivity, and nonverbal expressions of affect. Most items are rated on a 7-point scale (1 = normal, 7 = severe), with three items (sleep disturbance, appetite disturbance, and listless speech) rated on a 5-point scale. The interviewer uses clinical judgment to synthesize information from parent and child interviews to arrive at overall ratings that best describe the child for each item. The CDRS-R yields a total raw score range of 17 to 113, with CDRS $\geq$ 40 generally considered to be consistent with clinical depression. Although the CDRS-R was developed to assess depressive symptom severity in children age 6 to 12, it has been successfully employed in a number of studies that have included adolescents (e.g., Emslie et al., 1997). The CDRS-R has sound psychometrics, and its widespread use facilitates comparison across studies. Interrater reliability on summary scores

is excellent (r = .92; Poznanski & Mokros, 1996). The pediatric fluoxetine study by Emslie and his colleagues (1997) further demonstrated good interrater reliability for the CDRS-R with an intraclass correlation of $r = .95$. Test-retest reliability based on interviews at intake and 2 weeks later by different clinicians showed good stability (r = .80; Poznanski & Mokros, 1996). Internal consistency is considered adequate, with item-total correlations range from $r = .28$ for impaired schoolwork to $r = .78$ for depressed feelings (Poznanski & Mokros, 1996). The CDRS-R has good concurrent validity with the clinician-rated HAM-D, with score correlations of $r = .94$ for females and $r = .84$ for males (Brooks & Kutcher, 2001). Discriminant validity of the CDRS-R for detecting pediatric depression is reasonably good. It has been criticized for over-classifying depression, but the CDRS-R is generally employed to assess the severity of depressive symptoms rather than to establish diagnosis. Some have expressed concern that the CDRS-R—like the HAM-D for adults—overemphasizes somatic symptoms (Pavuluri & Birmaher, 2004) which tend to respond to antidepressants before there are detectable changes in mood. The CDRS-R interview requires approximately 30 to 45 minutes to administer. In the present study, the CDRS-R was employed to establish baseline symptom severity and to classify treatment outcome (remission or relapse).

CGI

The Clinical Global Impressions scale (Guy, 1976; NIMH, 1985) is a clinician-rated scale used to quantify the severity of psychiatric illness and the degree of improvement over time. The CGI scale was developed during the NIMH-PRB (Pharmacological Research Branch) collaborative schizophrenia studies (Schneider et al., 2002) and comprises three stand-alone subscales: (a) The CGI-Severity scale (CGI-S) reflects the clinician's judgment of the current overall level of mental illness as rated on a 7-point scale where 1 = normal, not ill, 2 = borderline mentally ill, 3 = mildly ill, 4 = moderately ill, 5 = markedly ill, 6 = severely ill, and 7 = among the most extremely ill. Ratings are necessarily anchored in the rater's clinical experience with the study population. (b) The CGI improvement scale (CGI-I) rates total improvement as compared with a defined baseline (e.g., study entry or start of treatment) on a similar 7-point scale where 1 = very much improved, 2 = much improved, 3 = minimally improved, 4 = no change, 5 = minimally worse, 6 = much worse, and 7 = very much worse. CGI-I $\leq$ 2 is often used in clinical trials to define an acceptable level of response to treatment. (c) The Efficacy Index (CGI-E) for drug studies is rarely used and was not employed in the present study. The CGI improvement scale provides the foundation for Quitkin pattern analysis. Its interrater reliability within the Emslie research group was verified in a previous study (Emslie et al., 1997). As a continuous variable (CGI-I ratings, 1 to 7), the intraclass correlation was .93.

When dichomotomized (CGI-I ≤ 2 = "improved" vs. CGI-I ≥ 3 = "not improved") the kappa coefficient was .95. Good reliability together with ease of use and intuitive scaling have made the CGI one of the most commonly used outcome measures in psychopharmacology trials, where it is often employed as a secondary measure alongside finer grained measures such as the HAM-D or CDRS-R. Importantly for the present study, the Quitkin method of pattern analysis uses sequential, dichotomized CGI-I ratings (improved vs. not improved) to characterize patient response to acute treatment with antidepressants.

Determination of Acute Response Patterns

Pattern analysis is based on improvements observed during acute (initial) drug treatment. In accord with the Quitkin protocol for adults, the present study employed pattern analysis to designate each pediatric patient as a likely true-drug or placebo-type responder to fluoxetine. Pattern determination required that a patient have at least seven of eight possible CGI improvement ratings for the acute phase (at weeks 1, 2, 3, 4, 6, 8, 10, 12). To facilitate pattern analysis, the CGI-I 7-point rating scale was dichotomized as described above (CGI-I ≤ 2 = improved vs. CGI-I ≥ 3 = not improved). Weeks for which a patient was improved are represented by the number 1, while unimproved weeks are represented by 0. This generated an eight-digit binary string for each patient to show the pattern of clinical improvement over the 12 weeks of acute fluoxetine

treatment (see Table A in the Appendix). Theoretically, this eight-digit string can consist of any combination of zeroes and ones, for example, 0010 1111. The space between the fourth and fifth digits renders the sequence more readable. The first (early) cluster of digits represents weekly ratings, and the second (later) cluster, biweekly ratings. *Early* response (improvement) as defined by the Quitkin and his colleagues (1984, 1987) is signified by a 1 appearing within the first 2 weeks, that is, in at least one of the first two digits of the binary string. Conversely, improvement is considered *delayed* if it does not occur before week 3, that is, if the first 1 appears in the third or later position in the sequence. Improvement is considered stable or *persistent* if, once achieved, it is maintained—that is, after the first 1 in the sequence, there are no subsequent 0s. If a 1 is later followed by a 0, improvement is considered fluctuating or *nonpersistent*. The present study follows the Quitkin adult protocol. Patients with a *delayed and persistent* response pattern are designated as likely 'true drug' responders, whereas those with *early and/or nonpersistent* patterns are termed 'placebo-type' responders. The following exemplify the method:

Sequence	Onset	Stability		Response pattern
0011 1111	delayed	persistent	→	true drug
0110 1110	early	nonpersistent	→	placebo type
0001 1101	delayed	nonpersistent	→	placebo type
0111 1111	early	persistent	→	placebo type
0000 0001	delayed	indeterminate	→	indeterminate (excluded)
0000 0000	—	—	→	nonresponsive (excluded)

Where a patient is missing one of the scheduled assessments and the two flanking values are equivalent (both = 1, or both = 0), the missing data point is assigned the equal value. If the missing data point is flanked by unequal values, the patient is excluded from the data analysis unless both 0 and 1 in the missing position result in the same pattern designation. Patients missing more than one rating were to be excluded, though no such cases were found. It should be noted that this pattern analysis paradigm allows for only three mutually exclusive and exhaustive drug response classifications: (a) true drug responder, (b) nonspecific, placebo-type responder, or (c) nonresponder. A fourth category, indeterminate, applies to patients who do not improve until the last treatment week (0000 0001), as persistence cannot be verified. There are no mixed pattern categories.

Statistical Analyses

Data Management

CDRR study data is computerized as part of the large affective disorders study database at the University of Texas Southwestern Medical Center (UTSW). CDRR data resides within the established and ongoing database for the Affective Disorders Child and Adolescent Research Group, managed by Dr. Carroll Hughes. Data are managed using a screen and menu guided application program for Microsoft Access 2000. Quality control procedures are in place to assure

complete and accurate data. All data are checked against the original documents, and dual entered into the database to ensure accuracy. Training sessions are conducted regularly to ensure high interrater reliability on key assessment measures, and frequently scheduled research meetings provide opportunity to resolve coding issues through discussion and consensus to enhance reliability. All clinical trainees work under close supervision.

Procedures are in effect to protect the confidentiality of research participants. Patients are identified by number in the computerized database, and are referred to only by number during research meetings. Patient research charts are kept locked at all times when not in use, and are available only to authorized study personnel. Study personnel are on call in the event of a patient emergency for which the patient's welfare requires access to research records.

For the present study, all data were analyzed using SPSS®, versions 11.5 and 16.0. Cox regression analysis was first conducted using SPSS®, then independently verified with SAS/STAT®, version 9.1.

Descriptive Statistics

Descriptive statistics are provided for relevant demographic and clinical variables. Means and standard deviations are reported for continuous variables, and frequencies and percentages provided for categorical variables. These statistics are reported for (a) the study group as a whole, as well being broken

down by (b) the two randomized continuation treatment groups, and (c) the two response pattern types previously described hypotheses. Group differences were assessed by independent *t* tests for continuous variables, and by chi square tests for dichotomized or categorical variables.

Hypothesis Testing of Between Group Differences

The four primary hypotheses (see Chapter 3) were tested by group comparisons on the dimension of relapse during the continuation phase. Comparisons were conducted between (a) patients randomized to fluoxetine versus placebo among those previously exhibiting true-drug acute response patterns, (b) patients randomized to fluoxetine versus placebo among those previously exhibiting placebo-type patterns, (c) patients with true-drug versus placebo-type patterns among those randomized to placebo, and (d) patients with true-drug versus placebo-type patterns among those randomized to fluoxetine. To test the four hypotheses—each addressing relapse for a pair of study groups—the primary analyses include the following:

1. Chi square tests to compare relapse rates/frequencies between groups.
2. Log-rank tests of Kaplan-Meier survival curves to assess the differences in time course to relapse between groups. Patients voluntarily leaving the study or lost to follow-up are censored at the point of last observation.

Prediction Modeling

Cox proportional hazards regression (Cox, 1972) is the procedure of choice for survival data, and is employed here to model patient risk for relapse.

Model specification. Considerable attention is devoted to covariate selection which is critical in Cox regression. Compared with ordinary linear regression, the semiparametric Cox procedure is more prone to instability when (a) important explanatory variables are omitted (referred to as "unobserved heterogeneity"), (b) multicollineary is present, (c) artificially dichotomized variables are employed, or (d) sample size is too small (Garson, n.d.; Harrell, 2001). Regarding sample size, the recommended number of covariates for Cox regression is based on the number of events (i.e., relapses) rather than number of subjects. The most commonly cited rule-of-thumb calls for a minimum of 10 to 15 events per variable (EPV; Concato, Peduzzi, Holford, & Feinstein, 1995; Harrell, Lee, Califf, Pryor, & Rosati, 1984; Peduzzi, Concato, Feinstein, & Holford, 1995). A recent simulation study by Vittinghoff and McCulloch (2007), however, suggests that this rule can be relaxed to 5 to 9 events per variable without undue compromise to model performance, and therefore the latter guideline is employed used here. Prior to covariate selection, 18 candidate variables were identified based on their perceived potential as predictors of relapse, their relevance to the stated hypotheses, and their ecological validity (i.e., common availability and usefulness) in research and clinical settings. In order to

narrow the field to the most promising predictors while minimizing problems with multicollinearity, a full correlation matrix was developed employing these 18 covariates plus the criterion variable, time to relapse, as well as relapse and full relapse. As both continuous and dichotomous variables are involved, correlational procedures variously include Pearson product-moment correlations, as well as phi, biserial, and point-biserial correlations. Variables included in the correlation matrix to inform covariate selection are the following:

1. Longitudinal pattern of response to acute fluoxetine treatment (nominal) = true drug or placebo-type (derived from sequenced CGI improvement ratings in the manner described above).
2. Randomized continuation treatment (nominal) = fluoxetine or placebo.
3. Age in years at study entry (numeric).
4. Age group at study entry (nominal) = child or adolescent. A patient of 11 years or younger is classified as a child, while a patient 12 or older is designated an adolescent. Regarding age group, at the time subjects were randomized for continuation treatment in the CDRR study, stratification by age was guided by an older convention, dividing patients into those 12 and younger (child) versus 13 and older (adolescent). The potential effect of this minor change is evaluated in descriptive statistics that reflect the adequacy of randomization.
5. Gender (nominal) = male or female.

6. Ethnicity (nominal) = African American, Caucasian, Hispanic, or other, recoded as Caucasian or other to reflect the preponderance of Caucasian participants.
7. Age in years at onset of first depressive episode (numeric).
8. Number of depressive episodes (numeric).
9. Current episode duration in weeks (numeric).
10. Depression chronicity (nominal) = chronic or nonchronic. The DSM chronic specific for mood disorders requires that MDD criteria be met continuously for 2 years, but this criterion may occur too infrequently in young populations to be useful. As such, for the present study, the episode duration requirement was reduced from 2 years to 48 weeks.
11. Neurovegetative symptom status (nominal) = positive or reverse. Neurovegetative status is determined from baseline K-SADS-PL data. Neurovegetative symptom status is considered 'reverse' for patients with both (a) hypersomnia and (b) increased appetite or weight gain; for all others, it is considered positive.
12. Comorbidity of MDD with another DSM Axis I diagnoses (nominal) = present or absent.
13. Dysthymia (nominal) = present or absent.
14. Family history of mood disorder in a first-order relative (nominal) = present or absent.

15. Pretreatment severity (numeric) = CDRS-R total score at acute baseline (CDRS-R_0).
16. Post-treatment severity (numeric) = CDRS-R total at randomization baseline (CDRS-R_{12}).
17. Post-treatment global severity (numeric) = CGI-S at randomization baseline (CGI-S_{12}).
18. Degree of symptom remission during acute treatment (numeric) = percentage drop in CDRS-R score, calculated as follows:

    $P_{CDRS\text{-}R} = 100 * (CDRS\text{-}R_0 - CDRS\text{-}R_{12}) / (CDRS\text{-}R - 17)$.

Although Cox regression employs time to relapse as the criterion variable, a total of three relapse outcome variables were included in the correlation analyses:

1. Relapse (nominal), defined as CDRS-R $\geq$ 40, or clinical deterioration suggestive or relapse (i.e., CRDR-R $\geq$ 28 accompanied by clinical judgment of significant risk for relapse) = present or absent.
2. Full relapse (nominal), defined as CDRS-R $\geq$ 40, preceded by 2 weeks of clinical deterioration = present or absent.
3. Time to relapse (numeric) = number of weeks post-randomization to withdrawal from study in the presence of relapse.

Proportional hazards assumption. Cox regression is a semi-parametric procedure that does not require specification of an underlying hazard function, but does rest on the important assumption that the hazard ratio for each of the covariates remains approximately constant over time. Prior to running the Cox procedure, each covariate was checked to ascertain that there were no significant time X covariate interactions regarding risk for relapse that would signal time dependence and therefore violation of the proportional hazards assumption.

The selected covariates were entered into a full fixed Cox regression model using simultaneously entry to avoid the well-document problems of automated stepwise modeling procedures (Derksen & Keselman, 1992; Harrell, 2002; Harrell et al., 1984). The model was first run using the main definition of relapse (CDRS-R $\geq$ 40, or clinical deterioration suggestive of relapse), then run again using the criterion for "full relapse" (CDRS-R $\geq$ 40, only).

Power Analysis

Sample size was generated and fixed during the CDRR study. Post-hoc power analysis is provided for key statistical procedures to inform the interpretation of results ($\alpha = .05$).

CHAPTER FIVE

RESULTS

Among 331 pediatric patients evaluated for the CDRR study, 168 patients met study criteria and were enrolled in the 12-week acute treatment phase of open-label fluoxetine (see Figure 1). Among those entering acute treatment, 49 patients either left early or failed to show adequate improvement, while 17 met response criteria but declined to continue beyond acute treatment. The remaining 102 patients entered the randomized continuation phase and thus constitute the subjects for the present study.

Descriptive Statistics

Demographics

Descriptive data for the 102 randomized patients (Table 1) are similar to those of the larger acute treatment group from which they emerged. At the start of the CDRR study, these patients ranged in age from 7 to 18 years, with a mean of 11.5 years (SD = 2.8). Among them, 44.1% were age 12 or older and therefore designated adolescents. Males comprised 63.7% of randomized subjects. In terms of ethnic composition, 70.6% were identified as Caucasian, 8.8% as African American, and 14.7% as Hispanic, with the remaining 5.9% from other racial/ethnic groups.

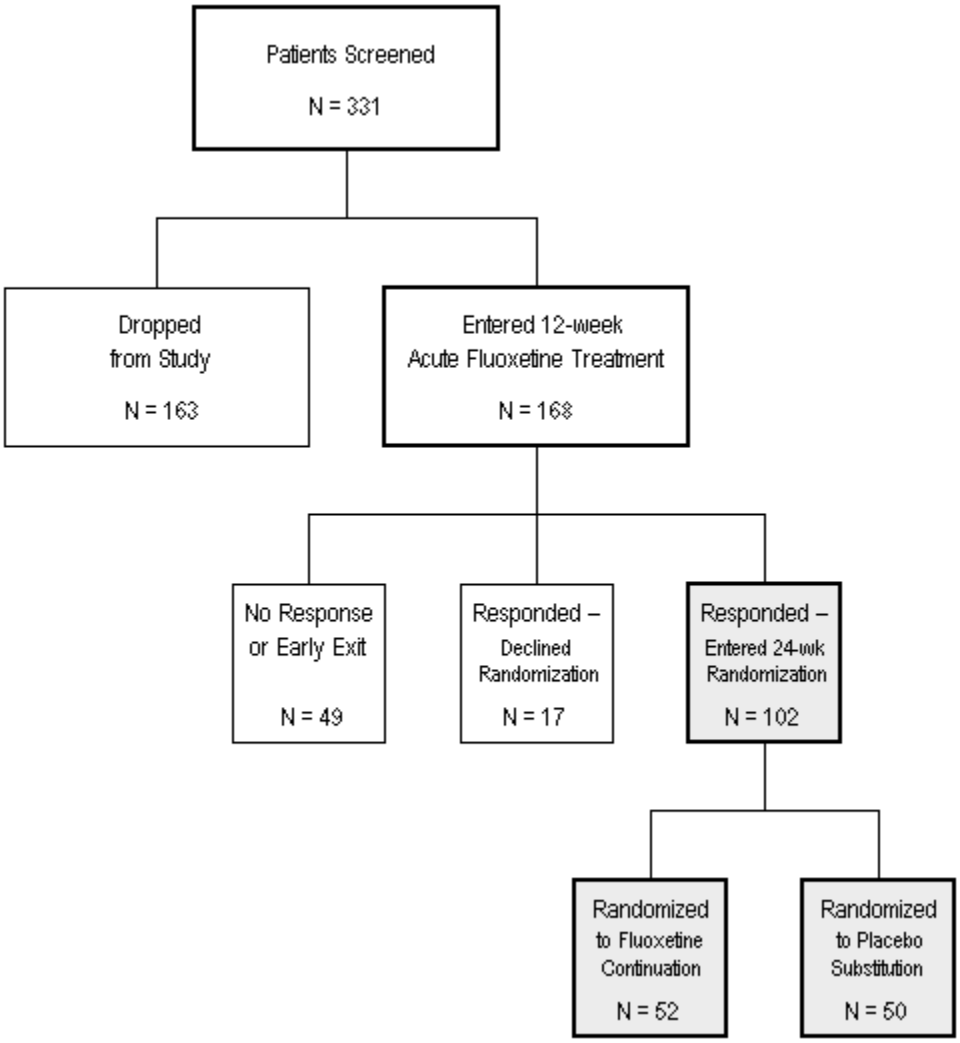

Figure 1. Participant flow chart

Table 1. *Demographic and Clinical Data*

	Continuation Phase N = 102	
	M (SD)	Frequency
Demographic		
Age (years)	11.5 (2.8)	
Age group		
Child (≤11 years)		57 (55.9%)
Adol (≥12 years)		45 (44.1%)
Gender		
Male		65 (63.7%)
Female		37 (36.3%)
Ethnicity		
Caucasian		72 (70.6%)
African American		9 (8.8%)
Hispanic		15 (14.7%)
Other		6 (5.9%)
Clinical		
Depression history		
No. of episodes	1.3 (0.5)	
Age at first onset	10.5 (2.8)	
Family history of mood D/O		76 (74.5%)
Current depressive episode		
First episode		74 (72.5%)
Age at onset	11.1 (2.7)	
Duration (weeks)	24.2 (19.9)	
Severity, CGI-S_0	4.8 (0.6)	
Severity, CDRS-R_0	57.8 (7.6)	
Comorbidity		
MDD only		27 (26.5%)
Dysthymia		34 (33.3%)
Anxiety disorder		26 (25.5%)
Behavior disorder[a]		45 (44.1%)
Other[b]		11 (10.8%)
Acute treatment response		
CGI-S_{12}	1.8 (0.6)	
CDRS-R_{12}	22.8 (4.2)	
CDRS-R drop (%)	85.5 (10.5)	

[a] Conduct disorder, oppositional defiant disorder, ADHD
[b] Enuresis, encopresis, trichotillomania, tic disorder, PTSD

Seventy-one percent (71.0%) of children (≤ 11 years) in the acute treatment phase advanced to randomized continuation treatment, while only 51.0% of adolescents did so, yielding a somewhat lower mean age for the subset of randomized participants (11.5 years, vs. 11.8 years in the acute phase). In terms of gender, just over half (37/71 = 52%) of the acute phase female patients entered the continuation phase, whereas two thirds (65/97 = 67%) of male patients were randomized. These demographic findings do not appear to be due to any differential effects of acute fluoxetine by age or gender, since comparisons between responders (n = 130) and nonresponders (n = 38) show no significant differences by either age ($t = 0.45$, $df = 166$, $p = .66$) or gender ($\chi^2 = 0.12$, $df = 1$, $p = .73$).

Clinical Features

Historical and baseline clinical features for patients who advanced to the continuation phase are summarized in Table 1 (above). Among the 102 randomized patients, the average age at onset for the first depressive episode was 10.5 years (SD = 2.8). Patients had experienced from 1 to 3 episodes, with a mean of 1.3 (SD = 0.50). Mean age of onset for the current episode was 11.1 years (SD = 2.7). For 72.5% of randomized patients, the index episode of depression was their first. Durations for this episode ranged from 1 month to 2.2 years, with a mean duration of just under 6 months (M = 24.2 weeks, SD = 19.9). On baseline evaluations using the Children's Depression Rating Scale−Revised

(CDRS-R) to assess depressive symptom severity, scores ranged from a low of 44 (*moderate*) up to 88 (*severe*) with a mean of 57.8 (SD = 7.6). On the 7-point Clinical Global Impressions–Severity (CGI-S) scale, ratings ranged from 4 to 6 (*moderately ill* to *severely ill*), with a mean of 4.8 (SD = 0.6; *markedly ill*). A positive family history of mood disorder in a first order relative was reported for three fourths (74.5%) of these patients.

Adequacy of Randomization

The patients in the two randomized continuation treatment groups (fluoxetine and placebo) were demographically and clinically similar (Table 2, below). All between group differences for the variables shown were statistically nonsignificant, suggesting that randomization was adequate. As reported by Emslie et al. (2008), however, the fluoxetine group had a higher baseline rate of comborbid anxiety disorders than the placebo group (36.0% and 15.4%, respectively; $\chi^2 = 5.70$, $df = 1$, $p = .02$).

Table 2. *Demographic and Clinical Data by Continuation Treatment*

	Fluoxetine Continuation N = 50		Placebo Substitution N = 52	
	M (SD)	Frequency	M (SD)	Frequency
Demographics				
Age (years)	11.2 (2.6)		11.8 (2.9)	
Age group				
Child (≤11 years)		29 (58.0%)		28 (53.8%)
Adol (≥12 years)		21 (42.0%)		24 (46.2%)
Gender				
Male		32 (64.0%)		33 (63.5%)
Female		18 (36.0%)		19 (36.5%)
Ethnicity				
Caucasian		34 (68.0%)		38 (73.1%)
African American		5 (10.0%)		4 (7.7%)
Hispanic		8 (16.0%)		7 (13.5%)
Other		3 (6.0%)		3 (5.8%)
Clinical				
Depression history				
No. of episodes	1.3 (0.5)		1.4 (0.6)	
Age at first onset	10.3 (2.7)		10.6 (2.9)	
Fam Hx mood D/O		35 (70.0%)		41 (78.8%)
Current episode				
First episode		37 (74.0%)		37 (71.2%)
Chronic (≥ 48 wks)		6 (12.0%)		8 (15.4%)
Reverse NVeg Sx		5 (10.0%)		5 (9.6%)
Age at onset	10.8 (2.5)		11.3 (2.9)	
Duration (weeks)	22.7 (18.6)		25.6 (21.1)	
Severity, CGI-S_0	4.8 (0.6)		4.9 (0.6)	
Severity, CDRS-R_0	57.5 (7.0)		58.0 (8.1)	
Acute fluox response				
True drug pattern		23 (46.0%)		27 (51.9%)
Placebo pattern		27 (54.0%)		24 (46.2%)
Indeterm pattern		0 (0.0%)		1 (1.9%)
CGI-S_{12}	1.8 (0.6)		1.7 (0.7)	
CDRS-R_{12}	23.3 (3.9)		22.4 (4.4)	
CDRS-R drop (%)	84.6 (9.4)		86.4 (11.4)	

Pattern Data

Binary pattern strings based on sequential CGI-I ratings were developed for each of the 137 patients who completed 12 weeks of acute fluoxetine treatment. As found by Quitkin et al. (1984, 1987), certain pattern strings were more common than others. Out of 256 (2^8) possible eight-digit binary sequences, only 24 different sequences were observed. Among the 102 patients advancing to the continuation phase, approximately half exhibited a true drug pattern of improvement (delayed and persistent improvement, 50/102 = 49.0%), and half showed a placebo pattern (early or nonpersistent, 51/102 = 50.0%); one patient's pattern was indeterminate (1/102 = 1.0%). The distributions of timing and persistence attributes for all patients completing acute treatment (n = 137) and for the subset who progressed to the randomization phase (n = 102) are provided in Table 3.

Table 3. *Frequency of Acute Response Pattern Attributes*

	Patients Completing Acute Treatment N = 137				Patients Continuing to Randomization Phase N = 102			
	Persist	Nonpers	Indeterm	Total	Persist	Nonpers	Indeterm	Total
Early	46	8	0	54	42	5	0	47
Delayed	63[b]	13	2	78	50[b]	4	1	55
No response	–	–	–	5	–	–	–	–

[a] Abbreviations denote Persistent, Nonpersistent, and Indeterminate
[b] True drug pattern

Patient Characteristics by Pattern

Comparisons of demographic and clinical characteristics of true drug versus placebo pattern responders are summarized in Table 4 (next page). Between group differences were not statistically significant except for the age variables. Those designated as placebo pattern responders were, on average, a year younger than true drug pattern responders (M = 10.9 years vs. 12.0 years, respectively; $t = 1.97$, $df = 99$, $p = .05$). It follows that placebo pattern responders were also younger at the time of onset of the index episode than true drug pattern responders (M = 10.5 vs. 11.7 years, respectively; $t = 2.26$, $df = 99$, $p = .03$).

Pediatric vs. Adults Response Patterns

Among those being treated with antidepressants, the adult studies of Quitkin et al. (1984, 1987) observed the same 50-50 split between true drug and placebo pattern responders as observed in the present study of younger patients. Pattern components, however—that is timing and stability of improvement—were apportioned differently for pediatric patients. Data comparing pediatric and adults response patterns are presented in Tables 5 and 6, below. Pediatric data in these tables include all responders in the present study with determinate patterns who completed the acute phase of fluoxetine treatment (n = 130). Adult data were pooled from the early studies by Quitkin et al (1984, 1987), which are the only published studies providing detailed breakdowns of pattern data; only those

Table 4. *Demographic and Clinical Data by Acute Response Pattern*

	True Drug Pattern N = 50		Placebo Pattern N = 51	
	M (SD)	Frequency	M (SD)	Frequency
Demographics				
Age (years)	12.0 (3.1)		10.9 (2.3)	
Age group				
Child (≤11 years)		25 (50.0%)		32 (62.7%)
Adol (≥12 years)		25 (50.0%)		19 (37.3%)
Gender				
Male		34 (68.0%)		31 (60.8%)
Female		16 (32.0%)		20 (39.2%)
Ethnicity				
Caucasian		39 (78.0%)		32 (62.7%)
African American		4 (8.0%)		5 (9.8%)
Hispanic		3 (6.0%)		12 (23.5%)
Other		4 (8.0%)		2 (3.9%)
Clinical				
Depression history				
No. of episodes	1.4 (0.6)		1.2 (0.5)	
Age at first onset	10.9 (3.1)		10.0 (2.4)	
Fam Hx mood D/O		36 (72.0%)		39 (76.5%)
Current episode				
First episode		32 (64.0%)		41 (80.4%)
Chronic (≥ 48 wks)		7 (14.0%)		6 (11.8%)
Reverse NVeg Sx		5 (10.0%)		5 (9.8%)
Age at onset	11.7 (3.1)		10.5 (2.2)	
Duration (weeks)	20.9 (15.8)		25.7 (19.9)	
Severity, CGI-S_0	4.9 (0.7)		4.8 (0.6)	
Severity, CDRS-R_0	58.0 (6.8)		56.9 (7.1)	
Acute fluox response				
CGI-S_{12}	1.8 (0.7)		1.7 (0.6)	
CDRS-R_{12}	23.4 (4.3)		22.2 (4.0)	
CDRS-R drop (%)	84.4 (10.5)		86.6 (10.5)	

assigned to acute treatment on active antidepressants are included here (n = 111). Table 5 shows the proportional distribution of patients in each cell of a 2 X 2 distribution of response timing and persistence for pediatric patients and adults.

Table 5. *Children vs. Adults: Proportional Distribution of Response Pattern Attributes*

	Pediatric Response Patterns[a] (from present study) N = 130[a]			Adult Response Patterns[b] (from Quitkin at al., 1984, 1987) N = 111[a]		
	Persistent	Nonpersistent	Total	Persistent	Nonpersistent	Total
Early	.354	.061	.415	.126	.243	.369
Delayed	.485[b]	.100	.585	.496[b]	.135	.631
Total	.839	.161	1.000	.622	.378	1.000

[a] Includes all responders with determinate patters of response to acute drug treatment
[b] True drug pattern

Table 6 (below) compares pattern frequencies in children versus adults, including frequencies of each of the two pattern components—improvement onset (timing) and stability (persistence). As noted above, for both adults and children. true drug and placebo response patterns were observed to occur with equal frequency (each approximately 50%). In terms of onset timing, children on antidepressant showed early improvement slightly more often than adults, but this difference was not statistically significant (42% vs. 37%, respectively; $\chi^2 = 0.53$, $df = 1$, $p = .47$). The difference between children and adults was more pronounced with regard to improvement stability. Nonpersistence occurred in only 16% of improving

pediatric patients, whereas it occurred in 38% of improving adults ($\chi^2 = 14.58$, $df = 1$, $p = .0001$). This finding suggests that once improvement is achieved in children, it tends to be more stable than in adults, at least during the short term of acute treatment.

Table 6. *Children vs. Adults: Frequency of Response Patterns and Component Attributes*

	Frequency (%)		Between Group Differences	
	Pediatric Patients[a] N = 130	Adult Patients[b] N = 111	Chi Square Statistic[c]	*p*
Response pattern				
True drug	63 (48.5%)	55 (49.5%)	0.03	.87
Placebo type	67 (51.5%)	56 (50.5%)		
Improvement onset				
Early	54 (41.5%)	41 (36.9%)	0.53	.47
Delayed	76 (58.5%)	70 (63.1%)		
Improvement stability				
Persistent	109 (83.8%)	69 (62.2%)	14.58**	.0001
Nonpersistent	21 (16.2%)	42 (37.8%)		

*p < .05, **p < .01
[a] From the present pediatric study; includes responders to acute fluoxetine treatment with determinate patterns
[b] From Quitkin et al. (1984, 1987); includes responders to antidepressant treatment with determinate patterns
[c] For chi square statistics, *df* = 1

Primary Hypotheses: Survival Analysis

Relapse Frequency

Although the principal analyses are Kaplan-Meier survival curves with log rank tests, an overview of the data is first presented in bar graph form (Figure 2). Patients who exhibited a true drug response pattern during acute treatment showed

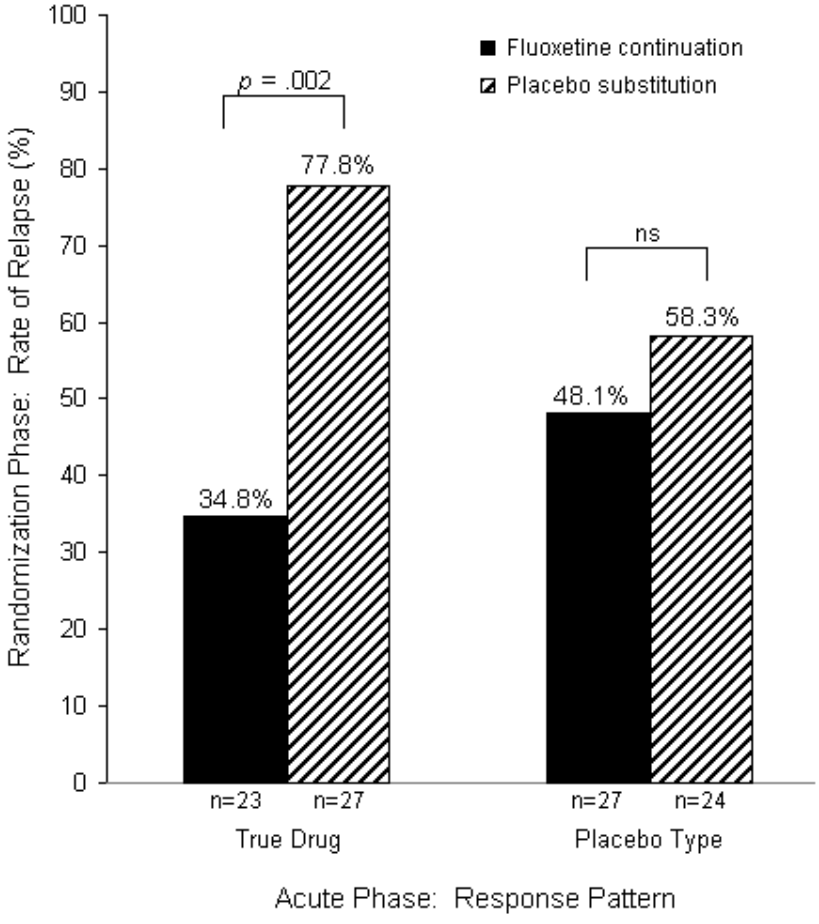

Figure 2. *Relapse frequency by pattern and treatment*

relapse rates of 34.8% (8/23) when continued on fluoxetine, versus 77.8% (21/27) when switched to placebo ($\chi^2 = 9.43$, $df = 1$, $p = .002$); the fluoxetine-placebo treatment effect (difference) was 43%. Placebo pattern responders exhibited a relapse rate of 48.1% (13/27) when continued on fluoxetine, compared with 58.3% (14/24) when switched to placebo ($\chi^2 = 0.53$, $df = 1$, $p = .47$); here the treatment difference was 10%. Looking only at patients randomized to placebo, the difference in relapse rates between true drug and placebo pattern responders did not reach significance ($\chi^2 = 2.23$, $df = 1$, $p = .14$). For patients continued on fluoxetine, relapse rates of true drug and placebo pattern responders were not significantly different ($\chi^2 = 0.91$, $df = 1$, $p = .34$). These findings are summarized in Table 7.

Table 7. *Relapse Frequency by Pattern and Treatment*

	Relapse Frequency (% who relapsed)	Betw Group Differences	
		Chi Square[a]	*p*
True drug pattern response to acute fluoxetine			
Randomized to fluoxetine (n = 23)	8 (34.8%)	9.43**	.002
Randomized to placebo (n = 27)	21 (77.8%)		
Placebo pattern response to acute fluoxetine			
Randomized to fluoxetine (n = 27)	13 (48.1%)	0.53	.47
Randomized to placebo (n = 24)	14 (58.3%)		
Randomization to placebo substitution			
Acute true drug pattern (n = 27)	21 (77.8%)	2.23	.14
Acute placebo pattern (n = 24)	14 (58.3%)		
Randomization to fluoxetine continuation			
Acute true drug pattern (n = 23)	8 (34.8%)	0.91	.34
Acute placebo pattern (n = 27)	13 (48.1%)		

*p < .05, **p < .01 [a] For chi square statistics, $df = 1$

The above results refer to the primary definition of relapse as described in Chapter 3 and Emslie et el. (2008). When similar data were generated to examine relapse frequency based on the stricter definition of relapse (i.e., "full relapse," CDRS-R $\geq$ 40 only), pattern effects were diminished. Patients who demonstrated a true drug pattern of response during acute treatment experienced full relapse at a rate of 21.7% (5/23) when continued on fluoxetine, versus 51.9% (14/27) when switched to placebo ($\chi^2 = 4.78$, $df = 1$, $p = .03$); the fluoxetine-placebo difference effect was 30.2%. Patients who showed a placebo pattern of response experienced full relapse at a rate of 22.2% (6/27) when continued on fluoxetine as compared with 45.8% (11/24) when switched to placebo ($\chi^2 = 3.19$, $df = 1$, $p = .07$); the difference effect was 23.6%. As pattern analysis predicts, the latter is statistically nonsignificant, though not persuasively.

Power analysis. As a retrospective study with a fixed n-size, the above 2 X 2 chi squares, each involving n = 50 or 51, were underpowered. According to Cohen (1988), achieving statistical power ($1 - \beta$) of .80 to detect a medium effect size of .30 (e.g., 60% relapse on placebo minus 30% on fluoxetine) at $\alpha = .05$ requires n = 90 for each analysis. At n $\approx$ 50, power was .56.

Survival Analysis

Statistics relating to Kaplan-Meier survival analysis for the four primary hypotheses are summarized in Table 8 at the end of this section.

Hypothesis 1. Patients with a true drug pattern of response to acute fluoxetine treatment will be more likely to relapse on placebo than fluoxetine during the randomized continuation treatment.

Among patients with true drug patterns, Kaplan-Meier analysis indicates that survival time without relapse is significantly greater in patients on fluoxetine than placebo (Figure 3; log rank score = 8.55, *df* = 1, *p* = .003). Survival estimates for mean time to relapse (based on both relapsed and censored patients) were 17.5 weeks on fluoxetine, compared with 9.6 weeks on placebo (P = 8.9 weeks).

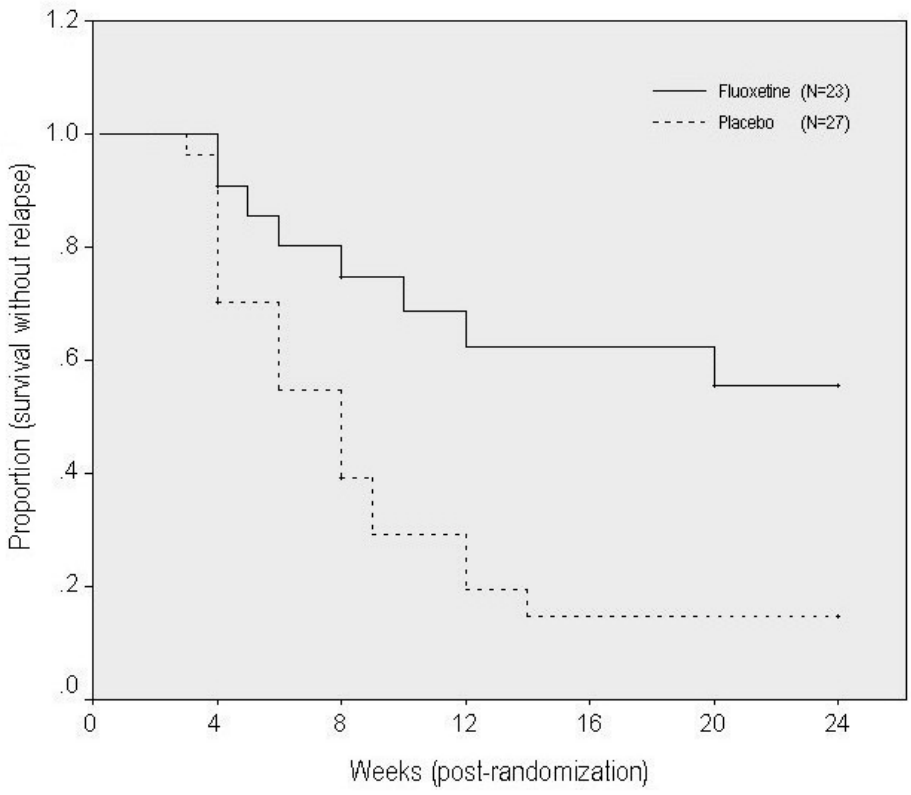

Figure 3. *True drug pattern responders: Kaplan-Meier survival curves of time to relapse on fluoxetine continuation vs. placebo substitution (log rank score = 8.55, df = 1, p = .003).*

Hypothesis 2. For patients with a nonspecific/placebo pattern of response during acute fluoxetine treatment, relapse during the continuation phase will *not* depend on continuation treatment (drug vs. placebo).

Among placebo pattern responders, Kaplan-Meier survival analysis showed no significant difference in relapse for those randomized to fluoxetine continuation versus placebo substitution (Figure 4; log rank score = 0.96, $df = 1$, $p = .33$). Survival estimates for mean time to relapse were 16.3 weeks on fluoxetine and 13.7 on placebo (P = 2.6 weeks).

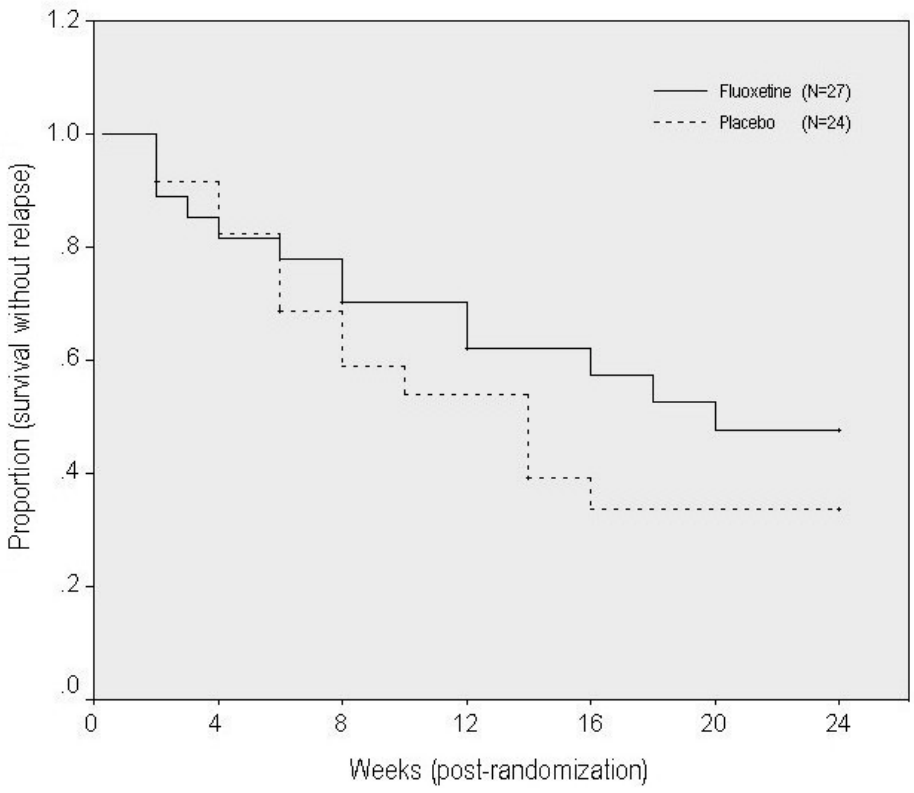

Figure 4. *Placebo pattern responders: Kaplan-Meier survival curves of time to relapse on fluoxetine continuation vs. placebo substitution (log rank score = 0.96, df = 1, p = .33).*

Hypothesis 3. For patients switched to placebo treatment during randomization continuation treatment, those who previously exhibited a true drug pattern of response to fluoxetine will relapse at a higher rate than those who previously exhibited a nonspecific/placebo pattern of response to fluoxetine).

For patients switched to placebo, Kaplan-Meier survival analysis of the difference in relapse for true drug versus placebo pattern responders showed a trend in the expected direction, although this difference did not reach statistical significance (Figure 5; log rank score = 2.92, $df = 1$, $p = .09$).

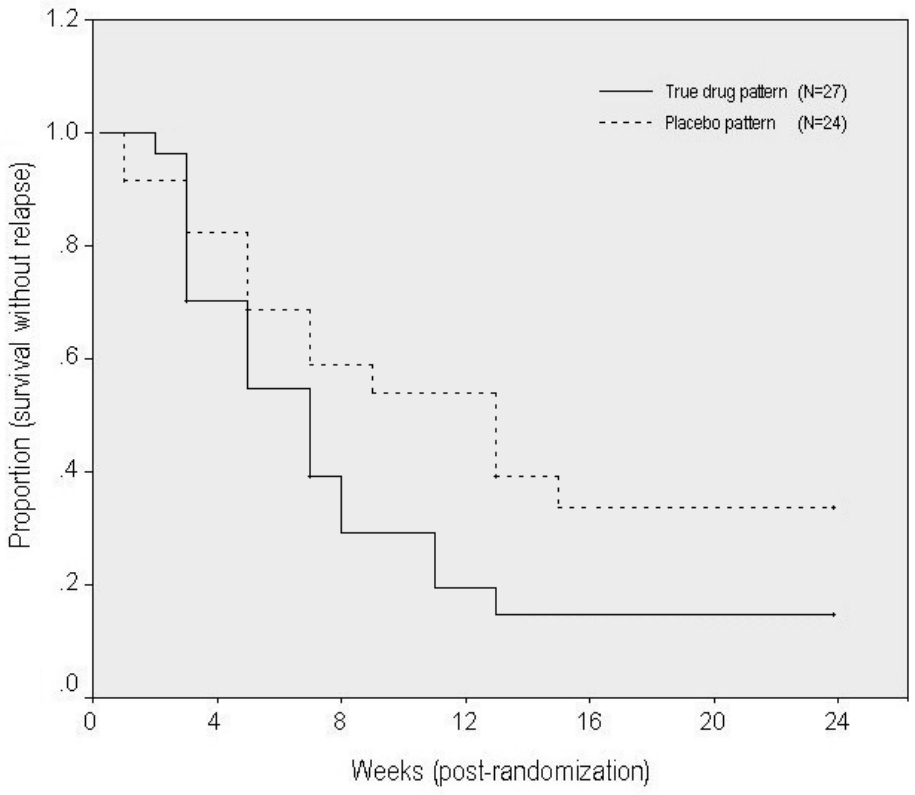

Figure 5. *Placebo substitution: Kaplan-Meier survival curves of time to relapse for true drug vs. placebo pattern responders (log rank score = 2.92, df = 1, p = .09).*

Hypothesis 4. For patients randomized to fluoxetine treatment during continuation, those who previously exhibited a true drug pattern of response will have a lower rate of relapse than those who previously exhibited a nonspecific/placebo pattern of response to fluoxetine).

Findings did not support the hypothesis. For patients randomized to fluoxetine continuation, Kaplan-Meier survival analysis showed no significant difference in relapse for true drug versus placebo pattern responders, who did equally well continuing on fluoxetine (Figure 6; log rank score = 0.29, $df = 1$, $p = .59$).

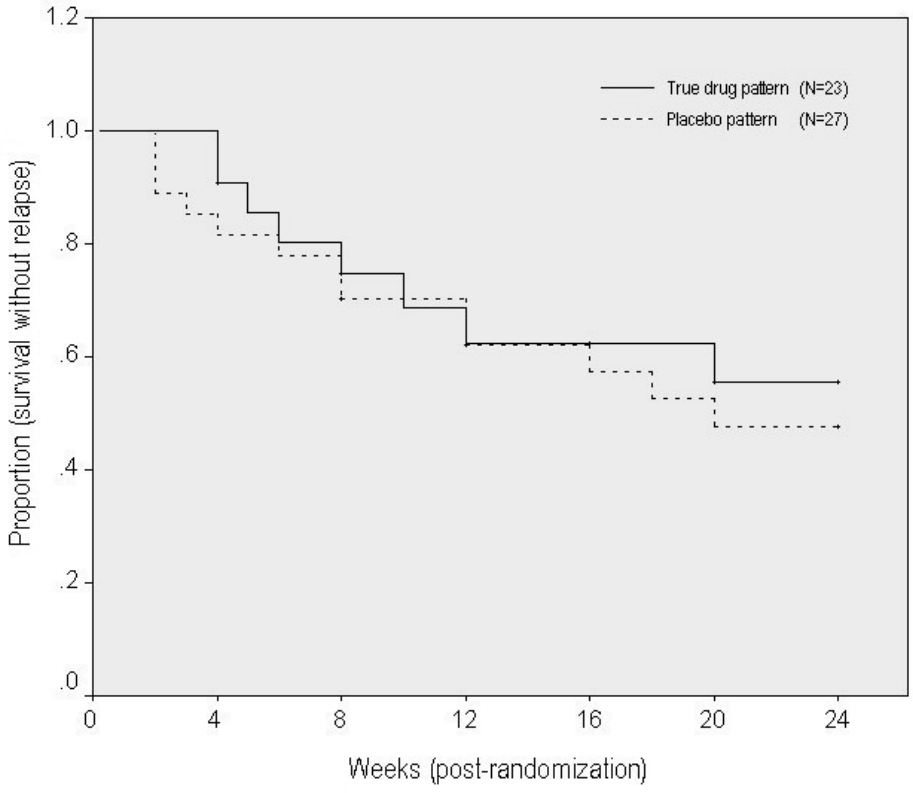

Figure 6. *Fluoxetine continuation: Kaplan-Meier survival curves of time to relapse for true drug vs. placebo pattern responders (log rank score = 0.29, df = 1, p = .59).*

Table 8. *Estimated Survival by Pattern and Treatment (Kaplan-Meier estimates)*

	Est. Time to Relapse (weeks)	Test of Equality for Survival Distributions	
	M (SE)	Log Rank Statistic[a]	*p*
Hypothesis 1: True drug pattern response			
Randomized to fluoxetine (n = 23)	17.5 (1.9)	8.55**	.003
Randomized to placebo (n = 27)	9.6 (1.4)		
Hypothesis 2: Placebo pattern response			
Randomized to fluoxetine (n = 27)	16.3 (1.7)	0.96[b]	.33
Randomized to placebo (n = 24)	13.7 (1.8)		
Hypothesis 3: Randomized to placebo			
Acute true drug pattern (n = 27)	9.6 (1.4)	2.92	.09
Acute placebo pattern (n = 24)	13.7 (1.8)		
Hypothesis 4: Randomized to fluoxetine			
Acute true drug pattern (n = 23)	17.5 (1.9)	0.29	.59
Acute placebo pattern (n = 27)	16.3 (1.7)		

*p < .05, **p < .01 [a] For log rank tests, df = 1 [b] Hypothesized no significant difference (none found)

Secondary Hypotheses: Regression Analysis

There were two purposes in using multivariate regression analysis to further examine relapse. First, Cox regression modeling is useful in identifying additional variables that might contribute to relapse and, unlike Kaplan-Meier survival analysis, both discrete and continuous variables may be investigated. Second, regression analysis holds other covariates constant in order that factors of interest can be "more purely estimated" (Harrell, 1996, p. 361).

Model Specification

According to the guidelines of Vittinghoff and McCulloch regarding the number of events per variable (EPV = 5 to 9; 2007), the 56 relapse events in the present study allow for a model of 6 to 11 covariates. Based on this, examination of the correlation data was aimed at selecting up to 10 covariates for the initial, full fixed model.

Correlational Findings

From among 1370 variables in the CDRR study, 18 clinical and demographic variables were identified for consideration as covariates for Cox regression analysis. The intent was to select covariates that were potentially good predictors of relapse and/or relevant to the stated hypotheses, and possessed ecologically validity for research and clinical settings. A correlation matrix of 18 potential predictors and 3 relapse variables was developed (Table 9), along with bivariate scatter plots of continuous variables (Figure 7). These procedures helped to identify promising predictors and expose potential problems with multicollinearity. While correlation analysis was undertaken specifically to inform covariate selection, it was also useful as an exploratory procedure to highlight findings of potential interest for post hoc testing or future studies. The multiple calculations in the absence of Bonferroni corrections create substantial risk of Type I error, and all such exploratory findings are reported cautiously.

Table 9. *Correlation Matrix[a] of Potential Regression Covariates (N = 102)*

	1 Tx	2 Patt	3 Age	4 AGp	5 Gen	6 Eth	7 AAO	8 No.	9 Dur	10 Chr	11 NVeg	12 Com	13 Dys	14 Fam	15 Sev_0	16 Sev_{12}	17 CGI_{12}	18 Chg	19 Rel	20 F Rel	21 Time
1 Treatment (continuation phase)	—	-.07	-.11	-.04	-.01	.06	-.06	-.06	-.07	-.05	-.01	.10	.01	-.10	-.03	.11	.12	-.09	-.27**	-.27**	.24*
2 Pattern of response		—	.19*	.13	-.08	-.17	.16	.19	-.13	.03	.00	.18	.13	-.05	.08	.15	.09	-.11	.05	.05	-.15
3 Age (years)			—	.82**	-.03	.06	.86**	.16	.05	-.04	.05	-.25*	.00	-.05	.19	.24*	.05	-.20*	.05	-.02	-.24*
4 Age group (child/adolescent)				—	.07	.08	.70**	.07	.16	.05	.03	-.23*	.00	-.12	.24*	.31**	.11	-.25*	.07	.05	-.30**
5 Gender					—	-.13	-.04	.02	.16	.11	-.16	-.01	-.01	.16	.16	-.04	-.11	.08	.18	.13	-.22*
6 Ethnicity (Cauc/non-Cauc)						—	.10	-.06	-.09	-.07	.00	.10	.18	.03	.09	.12	.11	-.14	-.12	.02	.00
7 Age at 1st MDD onset (years)							—	-.23*	-.09	-.12	.09	-.22*	.02	-.05	.16	.25*	.11	-.21*	-.01	-.02	-.21*
8 Number of episodes								—	-.14	-.13	.01	.06	.01	-.04	-.04	-.02	-.01	.02	.04	.03	.03
9 Current episode duration (wks)									—	.81	-.16	-.14	.04	.11	.29**	.12	-.04	-.05	.17	.14	-.20*
10 Chronicity (current ≥ 48 wks)										—	-.06	-.08	.20*	.10	.29**	-.03	-.12	.10	.07	.12	-.14
11 Neurovegetative Sx (pos/rev)											—	.03	.09	-.04	-.06	-.20*	-.07	.19	-.16	-.10	.06
12 Comorbidity												—	.42**	-.20*	-.12	.05	.06	-.05	-.04	.07	-.02
13 Dysthymia													—	-.16	.13	.08	.10	-.05	-.04	.22*	-.04
14 Family Hx of mood disorder														—	-.08	-.29**	-.33**	.29**	-.02	.01	-.06
15 Severity, baseline ($CDRS\text{-}R_0$)															—	.19	.17	.08	.07	.12	-.16
16 Severity, post-acute ($CDRS\text{-}R_{12}$)																—	.78**	-.95**	.14	.26**	-.20*
17 Severity, post-acute ($CGI\text{-}S_{12}$)																	—	-.71**	.08	.21*	-.07
18 Degree of change ($\Delta_{CDRS\text{-}R}$)																		—	-.12	-.22*	.18
19 Relapse																			—	.66**	-.58**
20 Full relapse																				—	-.43**
21 Time to relapse (weeks)																					—

*$p < .05$, **$p < .01$

[a]For correlations involving nominal variables, the valence (+/−) reflects arbitrary coding protocols, and the direction of association should not be interpreted from the table.

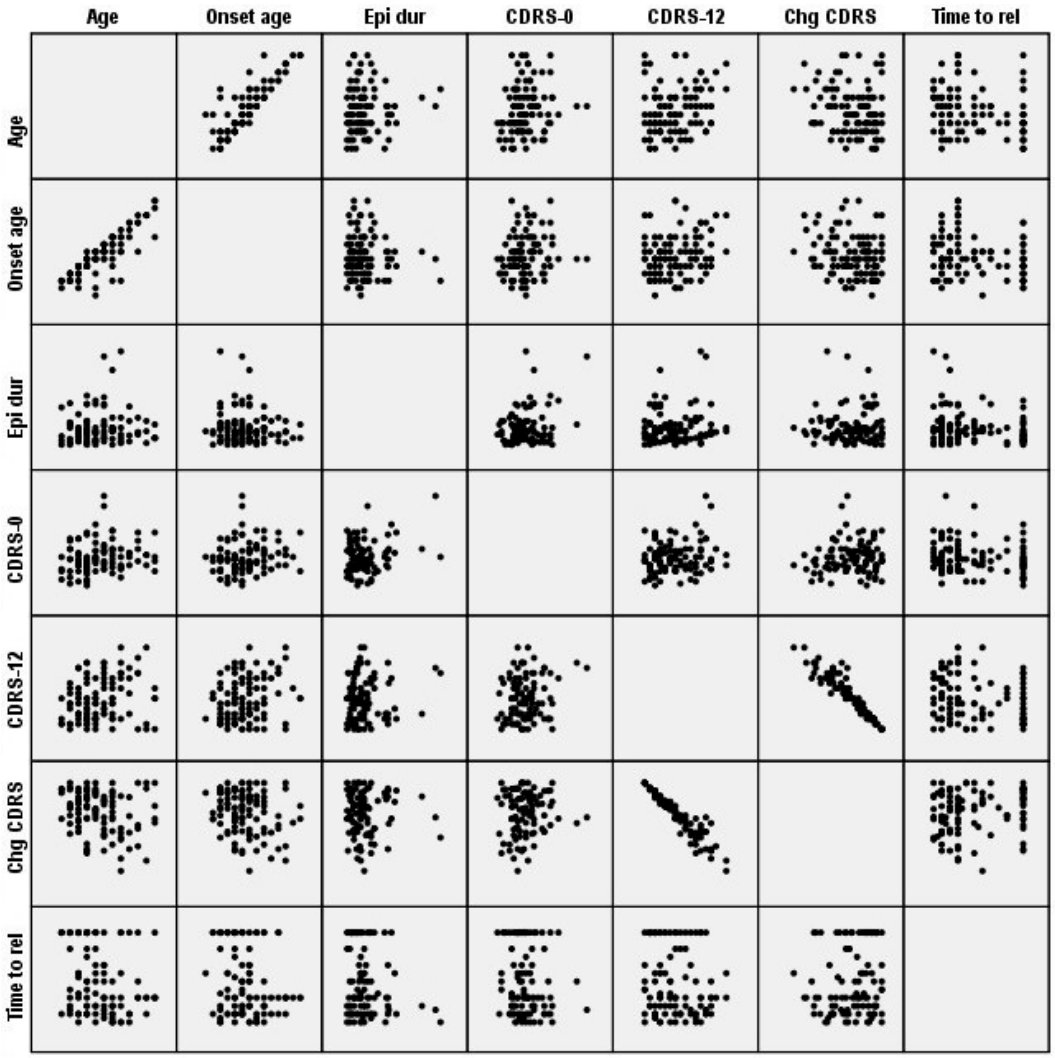

Figure 7. *Bivariate scatterplot matrix of continuous variables*

Relapse variables—relapse, full relapse, time to relapse. Although Cox regression uses time to relapse as the criterion variable, correlations were also developed for the nominal variables of relapse and full relapse. As expected, the three relapse variables showed significant correlations with one another—correlation of relapse with full relapse, $r_r = .66$; correlation of relapse with time to relapse, $r_b = -.58$; and correlation of full relapse with time to relapse, $r_b = -.43$. In the discussion that follows, these three variables—relapse, full relapse, and time to relapse—are collectively referred to as the "relapse variables."

Pattern and treatment. The interaction of acute response pattern (true drug vs. placebo type) and continuation treatment (fluoxetine vs. placebo) in predicting relapse are the focus of this study. Although treatment randomization was not stratified on acute response pattern, calculated correlations show no untoward association between them ($r_r = .07$). Consistent with prior findings (Emslie et al., 2008), continuation treatment (fluoxetine vs. placebo) was found to correlate with both relapse and full relapse ($r_b = .27$ in both cases), as well as time to relapse ($r_{pb} = .24$). Response pattern showed a modest correlation with age ($r = .19$), such that children were more likely than adolescents to show a placebo-type response to acute fluoxetine treatment.

Age and age group. Compared with other potential covariates, age group (child/adolescent) showed a surprising association with two separate outcome variables. (a) For the acute phase, age group showed a significant correlation with

percent change in total CDRS-R score ($r_b = .25$), such that childhood (vs. adolescence) was associated with a greater reduction in CDRS-R score during acute treatment. (b) In the continuation phase, age group was significantly correlated with time to relapse ($r_b = -.30$), with adolescents relapsing sooner. A post hoc test showed that, on average, children remain well (free of relapse) nearly 5 weeks longer than adolescents (means of 14.1 and 9.2 weeks, respectively; $t = 3.18$, $df = 100$, $p = .002$). Similar if somewhat lower correlations were found for age itself. As a continuous variable, age (in years) retains more information and was selected for inclusion in the regression analysis. Nonetheless, the strong showing for the age group variable supports the differentiation between children and adolescents as a developmentally meaningful one in terms of pediatric depression. Not surprisingly, age at initial onset was highly correlated with both age ($r = .86$) and age group ($r_b = .70$), and was excluded from the regression analysis to avoid problems of multicollinearity.

Gender. Gender did not show a significant correlation with either relapse or full relapse, but did correlate with time to relapse ($r_{pb} = .22$), with females relapsing sooner. Gender was included as a key covariate in regression analysis.

Ethnicity. The majority of CDRR participants in both the acute and continuation phases were Caucasian (75% and 71%, respectively). Following the convention of Emslie et al. (2008) for these patients, ethnic data was reduced to Caucasion and non-Caucasian; this dichotomized variable showed no relation to

time to relapse ($r_b = .00$). For the most part, correlations between specific ethnicities and various patient demographic and clinical characteristics were nonsignificant, except for the finding that Hispanic patients were somewhat more prone to placebo pattern responding than other ethnicities (12/15 = 80.0% for Hispanic patients vs. 39/87 = 44.8% for non-Hispanic patients; $\chi^2 = 6.14$, $df = 1$, $p = .013$). For the purposes of Cox regression analysis, ethnicity was represented by the dichotomized variable (Caucasian, non-Caucasian), although it was not expected to be a significant predictor of relapse.

Several clinical variables were considered for inclusion in the Cox model, including some that have shown predictive validity in adult studies of relapse.

Number of episodes. This variable showed no significant association with other variables, and showed virtually no correlation with relapse ($r_b = .04$), full relapse ($r_b = .04$), or time to relapse ($r = .03$). As no indications for using the number of depressive episodes in relapse prediction were found in the literature reviewed for this study, it was not included in regression analysis.

Episode duration/chronicity. In recent studies by the Quitkin group (especially McGrath et al., 2000, 2006), chronicity of depressive illness was found to be a robust predictor of relapse independent of treatment and was therefore proposed as a variable to be investigated in regression analysis for the present study. According to DSM criteria, the specifier designating a mood

disorder as "chronic" requires episode duration of 2 years of more, but this criterion is rarely met in younger patients (only 2 of 102 patients in the present study). In light of this, the operational definition for chronicity was adjusted downward to a duration of 48 weeks; this dichotomoized variable showed no significant correlation with the relapse variables. The continuous variable of episode duration (in weeks) showed a modest correlation with time to relapse ($r = -.20$) and was substituted for chronicity in regression analysis.

Neurovegetative status. Like chronicity, neurovegetative symptom status (positive or reverse), was selected for investigation based on prior findings by McGrath et al. (2000, 2006). In the present study, positive neurovegetative symptoms showed a modest negative correlation with post-acute treatment severity as measured by CDRS-R ($r = -.20$), suggesting that fluoxetine treatment may be somewhat more effective in these patients. Although neurovegetative status was not associated with any of the relapse variables, it was retained for the Cox model to determine if its effects of were differential by treatment.

Comorbidity. This variable represents the presence of at least one DSM Axis I diagnoses in addition to major depressive disorder. Comorbidity showed virtually no correlation with the relapse variables, and was therefore excluded from the Cox regression analysis. Comorbidity showed a modest negative correlation with the age variables, including age ($r = -.25$), age group ($r = -.23$), and age at depressive onset ($r = -.22$), suggesting that comorbidity may be

somewhat more prevalent in younger patients, although such results are reported with caution due to the risk of Type I error.

Dysthymia. The presence of dysthymia represented "double depression" in study patients and was expected to predispose to relapse. However, it showed virtually no correlation with either relapse or time to relapse, although there was a modest correlation with full relapse ($r = .22$). Dysthymia was not included as a covariate for the Cox regression analysis.

Family history. Family history of mood disorder in a first order relative showed no significant association with the relapse variables, but positive family history was moderately associated with response to acute fluoxetine treatment, including correlation with post-acute CGI-S ($r = -.33$), post-acute CDRS-R ($r = -.29$), and percent decline in CDRS-R ($r = .29$). These results suggest that family history of mood disorder may be associated with improved outcome during acute fluoxetine treatment. Although family history did not appear to be directly correlated with relapse during the continuation phase, it was included in the regression analysis in order to determine if the association between family history and response to acute fluoxetine extends to relapse prevention.

Depression severity—baseline and post acute. Acute baseline symptom severity (CDRS-R_0) was only weakly associated with continuation relapse. Since decisions regarding continuation treatment that might be usefully informed by pattern analysis occur *after* symptom remission has been achieved, the baseline

severity variable was excluded from regression analysis in favor of a more predictive post-acute measure. The available (and strongly collinear) post-acute response variables (CGI-S_{12}, CDRS-R_{12}, and $P_{CDRS\text{-}R}$), CDRS-R at week 12 (CDRS-R_{12}) showed the strongest correlation with relapse and was therefore selected for inclusion in the Cox regression analysis.

To summarize, based on a combination of study hypotheses, prior research (McGrath et al., 2000, 2006), and the above correlation analyses, the covariates selected for inclusion in the main Cox prediction model included response pattern, continuation treatment, age, gender, ethnicity, episode duration, neurovegetative status, family history of mood disorder, and post-acute treatment severity (CDRS-R_{12}). While prior adult studies (McGrath et al., 2000, 2006) have reportedly tested exploratory Cox models with all possible two- and three-way interactions, doing so for the nine covariates in the present study would have required an initial model of 129 terms. Given the limitations of the recommended events-per-variable ratio for the Cox procedure (minimum EPV between 5 and 9), the only interaction term considered in the initial model was pattern x treatment.

Cox Regression: Proportional Hazards Assumption

The proportional hazards assumption required by the Cox procedure was checked by examining time x covariate interactions. None were found to be significant, indicating that their respective hazard ratios are not time-dependent—

that is, the hazard ratios are sufficiently constant over time to meet the proportional hazards assumption.

Cox Regression: Treatment Only

Cox regression analysis is usually run as a multivariate procedure. Running the Cox procedure first with treatment alone, however, establishes a reference point for gauging the benefits of controlling for relevant covariates in subsequent multivariate models examining the effects of fluoxetine in reducing risk for relapse. With treatment alone as a single predictor in Cox regression (Table 10; previously reported in Emslie et al., 2008), patients randomized to placebo were at twice the risk for relapse as those on fluoxetine (hazard ratio = 2.10, 95% CI = 1.22–3.62; Wald $\chi^2 = 7.07$, $df = 1$, $p = .008$).

Table 10. *Cox Regression, Treatment as Sole Predictor (N = 101)*

	Parameter Estimate	SE	Wald χ^2	*p*	Hazard Ratio	Hazard Ratio 95% CI
Treatment	0.74	.28	7.07**	.008	2.10	1.22–3.62

*$p < .05$, **$p < .01$ For Wald χ^2, $df = 1$

In the multivariate models that follow, Cox regression effectively controls for other covariates, thus isolating the effects of each covariate of interest. It should be noted that the multivariate Cox procedure reported in Emslie et al. (2008, p. 463) employed a different and larger set of covariates (13 in all) that did not include response pattern, and results were reported only for the treatment parameter.

Cox Regression: Multivariate Models

Hypotheses 5/5a. Regression analysis will determinate that certain clinical and/or demographic variables contribute to the prediction of relapse—specifically including continuation treatment, as well as its interaction with acute response pattern.

The full fixed model included nine selected covariates plus one interaction: (a) continuation treatment, (b) acute response pattern, (c) age, (d) gender, (e) ethnicity, (f) episode duration (representing chronicity), (g) neurovegetative symptom status, (h) family history of mood disorder, (i) depression severity at randomization (CDRS-R_{12}), and (j) interaction of pattern and treatment. All covariates were entered simultaneously to retain the interpretive value of the resulting parameters and their *p*-values, which tend to be distorted by stepwise or other automated model building procedures (Derksen & Keselman, 1992; Harrell, 1996, 2002). Results for the initial full fixed model are shown in Table 11.

Table 11. *Cox Regression, Full Fixed Model: Demographic and Clinical Risks for Patients Randomly Assigned to Fluoxetine or Placebo After Acute Response to Fluoxetine*

	Parameter Estimate	SE	Wald χ^2	*p*	Hazard Ratio	Hazard Ratio 95% CI
Treatment	1.39	.45	9.51**	.002	4.02	1.66−9.75
Response pattern	0.49	.47	1.05	.31	1.63	0.64−4.12
Age	0.01	.05	0.08	.78	1.02	0.92−1.12
Gender	0.83	.30	7.59**	.006	2.30	1.27−4.17
Ethnicity	0.12	.34	0.13	.72	1.13	0.58−2.21
Episode duration (wks)	0.01	.01	2.36	.13	1.01	1.00−1.03
Neuroveg Sx status	−0.08	.42	0.04	.84	0.92	0.41−2.09
Family Hx, mood D/O	−0.14	.33	0.19	.66	0.87	0.46−1.65
Severity at randomiz[a]	0.06	.04	2.34	.13	1.06	0.99−1.13
Treatment x pattern	−1.08	.60	3.29	.07	0.34	0.11−1.09

*p < .05, **$p < .01$ For each Wald χ^2, $df = 1$ [a] CDRS-R_{12}

While computer-run stepwise procedures for model building have been strongly criticized in the medical statistics literature as providing unstable, inaccurate results (Derksen & Keselman, 1992; Harrell, 1996, 2001), Garson (n.d.), advocates judicious "model trimming" by investigators (rather than computer algorithms) to eliminate highly nonsignificant variables and achieve a more parsimonious model. Accordingly, the Cox analysis was rerun without the nonsignificant covariates of age, ethnicity, neurovegetative status, or family history—all of which showed *p*-values > .60 in the full fixed model. Results for the trimmed model are shown in Table 12. The six remaining covariates held steady (same values as in the full fixed model), thus providing a measure validation for both models.

Table 12. *Cox Regression, Trimmed Model*

	Parameter Estimate	SE	Wald χ^2	*p*	Hazard Ratio	Hazard Ratio 95% CI
Treatment	1.40	.44	9.98**	.002	4.04	1.70−9.61
Response pattern	0.43	.46	0.88	.35	1.53	0.63−3.75
Gender	0.83	.29	8.20**	.004	2.29	1.30−4.04
Episode duration (wks)	0.01	.01	3.02	.08	1.01	1.00−1.03
Severity at randomiz.[a]	0.05	.03	2.32	.13	1.05	0.99−1.11
Treatment x pattern	−1.08	.60	3.28	.07	0.34	0.11−1.09

*p < .05, **$p < .01$ For each Wald χ^2, $df = 1$ [a] CDRS -R_{12}

In order to comment on the full range of variables, the numeric results that follow refer to the full fixed model results in Table 9.

Treatment. The Cox regression findings regarding the treatment effect of fluoxetine versus placebo in reducing the risk for relapse during continuation treatment strongly reinforce the CDRR findings reported in Emslie et al. (2008) and parallel the findings for adult fluoxetine continuation studies (Geddes et al., 2003; McGrath et al., 2000, 2006; Montgomery et al., 1988). When taking account of (controlling for) covariate effects, the risk of relapse while on placebo is four times that of fluoxetine (hazard ratio = 4.02, 95% CI = 1.66–9.75; Wald $\chi^2 = 9.51$, $df = 1$, $p = .002$).

Pattern. As expected, response pattern itself was not predictive of risk for relapse. Interaction of pattern x treatment fell just short of significance ($p = .07$).

Demographics. Neither age nor ethnicity were found to be predictive of risk for relapse, while female gender, was strongly so, independent of treatment or other variables (hazard ratio = 2.30, 95% CI = 1.27–4.17; Wald $\chi^2 = 7.59$, $df = 1$, $p = .006$). To ascertain that this result was not merely an anomaly stemming from inclusion of the pattern covariate, the Cox model was rerun without the pattern variable, and gender remained a significant predictor of risk for relapse (hazard ratio = 2.00, 95% CI = 1.11–3.57; Wald $\chi^2 = 5.39$, $df = 1$, $p = .02$). To further investigate this unexpected finding, a separate post hoc Kaplan-Meier survival

analysis was conducted to compare time to relapse for males versus females (see Table 13). Survival estimates of median time to relapse for males and females were 16 weeks and 8 weeks, respectively (log rank test = 5.36, $df = 1$, $p = .02$), indicating that females tend to relapse sooner than males, independent of treatment.

Table 13. *Estimated Survival by Gender (Kaplan-Meier estimates)*

	N	Survival Estimate of Time to Relapse (weeks)	S.E.	95% C.I.	Log Rank Statistic[a]	*p*
Male	65	Mean[b] = 15.6	1.1	13.4 − 17.7	5.36*	.02
		Median = 16.0	3.7	8.7 − 23.3		
Female	37	Mean[b] = 11.3	1.4	8.6 − 14.1		
		Median = 8.0	1.8	4.4 − 11.6		

*p < .05
[a] Test statistic for equality of survival distributions between groups
[b] Mean limited to 24 weeks (end of study)

Females also relapsed with greater frequency than males (67.6% and 49.2%, although this latter difference did not reach significance ($\chi^2 = 3.22$, $df = 1$, $p = .07$). Supplementary Cox analysis revealed no significant treatment x gender interaction, but when Cox and Kaplan-Meier survival procedures were run separately by gender—first for males, then females—fluoxetine treatment effects were somewhat better for males than females across the board.

Other covariates in the Cox regression analysis did not independently predict time to relapse.

Hypothesis 5b. Neurovegetative status will interact with continuation treatment to enhance relapse prediction. Specifically, positive neurovegetative symptoms (insomnia, loss of appetite) will increase the effect of fluoxetine in preventing relapse. For those with reverse symptoms (hypersomnia, weight gain), fluoxetine will confer no more benefit than placebo in preventing relapse.

Due to the absence of a main effect for neurovegetative symptom status in the above models and the low incidence of reverse neurovegetative symptoms (10 of 102 patients = 9.8%), the interaction of neurovegetative status and continuation treatment was not evaluated in the Cox procedure. Consistent with the hypothesis, the 10 patients with reverse symptoms showed no fluoxetine treatment effect whatever, relapsing at the same high rate whether randomized to fluoxetine or placebo (4/5 = 80% for both), but these numbers are too small to be meaningful.

Hypothesis 5c. Chronicity of illness will predict outcome (relapse vs. remission) independent of treatment. As a corollary, pattern analysis will be more effective in predicting differential response to fluoxetine continuation versus placebo substitution in the nonchronically depressed patients.

Contrary to consistent findings by McGrath et al. (2000, 2006) for adults, chronicity of illness—represented here by episode duration—was not a significant predictor of relapse (hazard ratio = 1.01, 95% CI = 1.00–1.03; Wald $\chi^2 = 2.36$, $df = 1$, $p = .13$). To examine the corollary hypothesis, the dichomotized chronicity variable was examined, with 'chronic' defined as episode duration of 48 weeks or more. Relapse frequencies for the 88 nonchronically depressed patients (Table 14) show an enhanced drug-placebo treatment effect for true drug

pattern responders ($\chi^2 = 7.95$, $df = 1$, $p = .004$) as compared with placebo pattern responders ($\chi^2 = 0.54$, $df = 1$, $p = .46$). Among the 14 chronically depressed patients (Table 15), cell sizes are too small to draw any meaningful conclusions, although contrary to the hypothesis, the numbers are consistent with an enhanced treatment effect for true drug pattern responders here, too.

Table 14. *Nonchronically Depressed*[a]*: Relapse Frequency by Pattern and Treatment*

	Relapse	Between Group Differences	
	Frequency (%)	Chi Square Statistic	*p*
True drug pattern responders			
Randomized to fluoxetine (n = 19)	7 (36.8%)	7.95**	.004
Randomized to placebo (n = 24)	19 (79.2%)		
Placebo pattern responders			
Randomized to fluoxetine (n = 25)	11 (44.0%)	0.54	.46
Randomized to placebo (n = 20)	11 (55.0%)		

*p < .05, **p < .01 [a] Episode duration <48 week (n = 88)

Table 15. *Chronically Depressed*[a]*: Relapse Frequency by Pattern and Treatment*

	Relapse	Between Group Diff
	Frequency (%)	Fisher Exact Test *p*-value
True drug pattern responders		
Randomized to fluoxetine (n = 4)	1 (25.0%)	.34
Randomized to placebo (n = 3)	2 (66.7%)	
Placebo pattern responders		
Randomized to fluoxetine (n = 2)	2 (100.0%)	.67
Randomized to placebo (n = 4)	3 (75.0%)	

*p < .05, **p < .01 [a] Episode duration $\geq$48 week (n = 13)

Cox Regression: Multivariate Model for Full Relapse

The CDRR study protocol provided for the early withdrawal of patients exhibiting clinical deterioration short of full relapse. This is sound ethical practice, though it effectively truncates data collection before the survival experiment has run its course. For this reason, data analysis in the present study has emphasized the primary (less strict) definition of relapse that incorporates clinical deterioration, but certain key analyses were run both ways, first using the primary (less strict) definition for relapse, then using full relapse. Results for running the Cox regression model with *full relapse* as the criterion event are shown in Table 16. Continuation treatment and gender again yielded significant risk ratios, and severity at randomization (CDRS-R_{12}) also emerged as a strong predictor. The interaction of treatment X pattern was nonsignificant in this model.

Table 16. *Cox Regression, Full Fixed Model: Demographic and Clinical Risks for Full Relapse for Patients Randomly Assigned to Continuation Treatment on Fluoxetine or Placebo*

	Parameter Estimate	SE	Wald χ^2	*p*	Hazard Ratio	Hazard Ratio 95% CI
Treatment	1.43	.57	6.37*	.012	4.19	1.38−12.73
Response pattern	−0.03	.64	.00	.96	0.97	0.28−3.37
Age	−0.07	.07	1.00	.32	0.93	0.81−1.07
Gender	1.05	.39	7.17**	.007	2.85	1.32−6.13
Ethnicity	−0.21	.41	0.27	.60	0.81	0.36−1.80
Episode duration (wks)	0.02	.01	3.57	.06	1.02	1.00−1.04
Neuroveg Sx status	−0.43	.53	0.65	.42	0.65	0.23−1.84
Family Hx, mood D/O	−0.39	.44	0.81	.37	0.68	0.29−1.59
Severity at randomiz.[a]	0.14	.05	8.64**	.003	1.15	1.05−1.26
Treatment X pattern	−0.45	.76	0.34	.56	0.64	0.14−2.86

*p < .05, **p < .01 For each Wald χ^2, *df* = 1 [a] CDRS-R_{12}

Exploratory Hypotheses

Hypothesis 6: Differential Dropout

Acute response patterns will be associated with differential dropout rates on fluoxetine continuation versus placebo substitution. Specifically, true drug responders will drop out sooner and more frequently when randomized to placebo substitution (vs. fluoxetine). Conversely, placebo-type responders will not show differential dropout rates for fluoxetine and placebo.

During the continuation phase, one in six patients dropped out voluntarily ("patient or parent decision"), and the distribution was roughly equal across the four pattern x treatment groups (Table 17). Time to voluntary withdrawal during the continuation phase was approximately 7 weeks across the four groups, with the exception of the five placebo pattern responders randomized to fluoxetine who remained in the study an average of 11.4 weeks before withdrawing.

Table 17. *Voluntary Dropout by Pattern and Treatment*

	Frequency		Time (weeks) to Dropout M (SD)	
	True Drug Pattern	Placebo Pattern	True Drug Pattern	Placebo Pattern
Fluoxetine	4/23 (17.4%)	5/27 (18.5%)	7.5 (6.2)	11.4 (3.4)
Placebo	5/27 (18.5%)	3/24 (12.5%)	7.0 (3.6)	7.3 (6.1)

The analysis was expanded to examine pattern effects with regard to early exiting overall, including relapse, voluntary withdrawal, and other reasons. Because relapsing patients were removed early in the interests of their welfare,

relapse and exit week often coincided, and therefore the findings (Table 18) are redundant with and not qualitatively different from those in Hypotheses 1 and 2. True drug pattern responders on placebo tended to exit more frequently and significantly sooner than those on fluoxetine, though the latter finding fell short of significance Among placebo pattern responders, neither frequency nor timing of early exit was differential by treatment. Examination of the proportional distribution of early exiting reasons (Table 19) was also unenlightening.

Table 18. *Early Exit by Pattern and Treatment*

	True Drug Pattern				Placebo Pattern			
	Fluoxetine N = 23	Placebo N = 27	Betw Gp Differ	*p*	Fluoxetine N = 27	Placebo N = 24	Betw Gp Differ	*p*
Frequency (%) early dropout	16 (69.6%)	24 (88.9%)	$\chi^2 = 2.90$	.09	17 (63.0%)	18 (75.0%)	$\chi^2 = 0.86$	.34
Weeks to exit M (SD)	13.2 (8.7)	8.7 (6.2)	$t = 2.11$*	.04	14.6 (8.6)	11.8 (8.3)	$t = 1.20$	.24

*p < .05

Table 19. *Reasons for Early Exit by Pattern and Treatment*

	True Drug Pattern		Placebo Pattern	
	Fluoxetine N = 23	Placebo N = 27	Fluoxetine N = 27	Placebo N = 24
Relapse (clin deter)	3 (13.0%)	5 (18.5%)	5 (18.5%)	3 (12.5%)
Full relapse	5 (21.7%)	13 (48.1%)	6 (22.2%)	11 (45.8%)
Other[a]	8 (34.8%)	6 (22.2%)	6 (22.2%)	4 (16.7%)

[a] Parent decision, adverse event, noncompliance, protocol violation, or lost to follow-up

Hypothesis 7: Patient Characteristics by Pattern

True drug and placebo pattern responders will differ on some demographic and clinical attributes. (This was an exploratory hypothesis to determine differential characteristics, if any, between true drug vs. placebo pattern responders.)

Findings regarding the demographic and clinical attributes for the two patterns types are discussed under Descriptive Statistics, above. The main finding was that placebo pattern responders were over a year younger, on average, than true drug pattern responders (M = 10.9 years vs. 12.0 years, respectively; $t = 1.97$, $df = 99$, $p = .05$).

Hypothesis 8: Abrupt Improvement

Abrupt improvement (e.g., a jump in global improvement rating from CGI-I = 4 to CGI-I = 1 or 2 in consecutive weeks/assessments) during the acute phase will tend to occur within the first 2 weeks, and if so, not to persist. (As such, early abrupt improvement would provide an added marker for placebo-type responding.)

Among 102 study patients in the present study, 19 (18.6%) showed abrupt improvement during acute treatment. Consistent with the hypothesis, all such improvements occurred during the first 2 weeks of treatment—13 in the first week, 6 in the second—classifying these patients as placebo pattern responders. Contrary to expectations, only 1 of 19 abrupt responders showed nonpersistence. Among these abrupt responders, relapse rates were 33.3% (4/12) for those randomized to fluoxetine, and 57.1% (4/7) for those on placebo. Fisher's exact

test for small numbers showed this observed difference to be nonsignificant (1-tailed, p = .53).

Hypothesis 9: Pattern Duration

A six-digit acute response pattern based on the first 8 weeks of acute treatment (weeks 1, 2, 3, 4, 6, 8) will relate to outcome as effectively as the eight-digit pattern reflecting the full 12 weeks of treatment (weeks 1, 2, 3, 4, 6, 8, 10, 12).

When existing binary pattern strings were truncated at six digits, nine patients were lost to analysis: Four patients with the true drug pattern 0000 0011 converted to nonresponder status (0000 00), while four patients with the true drug pattern 0000 0111 and one patient with the placebo pattern 0000 0101 converted to an indeterminant pattern (0000 01). For the remaining 92 patients, the six-digit pattern strings (8 weeks) demonstrated an enhanced treatment effect for true drug versus placebo pattern responders that was comparable to that found for the eight-digit strings (12 weeks), suggesting that a shorter period of monitoring to establish acute response patterns may be equally effective (see Figure 8 compared with Figure 2).

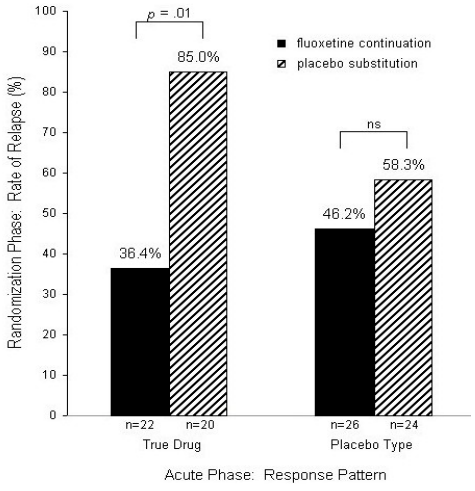

Figure 8. *Relapse frequency by pattern and treatment for 6-digit patterns (8 weeks)*

CHAPTER SIX

DISCUSSION

This is the first randomized, placebo-controlled study of Quitkin pattern analysis in a pediatric population. Widespread placebo effects associated with antidepressant treatment confound the results of drug trials and complicate clinical decisions regarding continuation treatment for relapse prevention. In adults, the Quitkin pattern paradigm has been shown to differentiate true drug benefits from placebo effects during acute antidepressant treatment. The goals for the present study were (a) to explore the validity of pattern analysis as applied to depressed children and adolescents, and (b) to examine the role of pattern analysis in the larger, multivariate context of modeling relapse prediction.

The data presented here were collected as part of the two-phase CDRR continuation study of fluoxetine for the prevention of relapse in pediatric depression (Emslie et al., 2008). Response data gathered during the 12-week acute treatment phase on open-label fluoxetine were used to designate patients as either true drug or placebo pattern responders in the manner of Quitkin et al. (1984, 1987). After acute treatment, 102 responders were randomized to 24 weeks of continuation treatment on fluoxetine or placebo and monitored for relapse. The interpretation and relevance of the findings reported above in Chapter 5 are considered below.

Pattern Analysis

The early studies of Quitkin et al. (1984, 1987) provide the only extant documentation of adult pattern distributions in response to acute antidepressant treatment, and the present study is the first to examine response patterns in pediatric patients.

Age. The finding that placebo pattern responders are, on average, a year younger than true drug responders (10.9 vs. 12.0 years, respectively) is consistent with prior studies and likely reflects developmental differences in young children—that is, greater dependence, compliance, and suggestibility, and more naïve expectations regarding medicine. An article by Bridge et al. (2007) reported that for 13 placebo-controlled drug trials for pediatric depression (total N = 1,552), children showed a 58% placebo response rate as compared with 49% for adolescents. This highlights the difficulty in establishing antidepressant efficacy in children, and in turn, the potential usefulness in being able to identify true drug responders in drug trials involving the youngest patients.

Treatment effects compared by pattern. The present study found that the fluoxetine-placebo treatment effect was considerably larger among true drug responders than among placebo pattern responders (Figure 2), thus supporting pattern analysis as a means to differentiate true drug benefits from placebo effects in young patient. As a point of reference, when Emslie et al. (2008) analyzed the

CDRR data in the absence of pattern information, they found lower relapse rates among patients randomized to fluoxetine (42.0%) compared with those switched to placebo (69.2%), for a treatment effect (difference) of 27.2%. In the present study, this drug-placebo treatment effect difference was enhanced to 43.0% (up from 27.2%) in true drug responders, and reduced to 10.2% (n.s.) among placebo pattern responders—that is, the observed treatment effect size for true drug responders was four times that of placebo pattern responders. Kaplan-Meier survival analysis produced similar results (Figure 3 vs. Figure 4). These findings parallel those of adult continuation studies (Nierenberg, 2004; Stewart et al., 1998), and indicate that Quitkin pattern classification captures key features of true fluoxetine response in the pediatric population, and thereby identifies those most likely to benefit from continuing medication to prevent relapse.

Pattern effects compared by treatment. Examining data from within each of the two continuation treatment groups, however, did not show the hypothesized pattern effects to be significant.

Among patients randomized to placebo, true drug responders did somewhat worse than placebo pattern responders, but these results were not statistically significant in terms of either relapse frequency ($p_{\chi^2} = .14$) or Kaplan-Meier survival estimates of time to relapse ($p_{\text{log rank}} = .09$). The hypothesized difference is, however, visible on the survival curves (Figure 5). Here the failure

to reach significance is likely due underpowered statistical analyses (e.g., for the chi squares reported here, power to detect a medium effect at $\alpha = .05$ was .56, which is well short of the usual recommendation that power = .80; Cohen, 1988). The heterogeneous nature of the placebo pattern designation (considered below) also may have contributed to nonsignificance.

For those randomized to continue on fluoxetine, the hypothesized pattern differences simply did not emerge (Figure 6)—that is, placebo pattern responders, who were not expected to benefit from fluoxetine, fared almost as well as true drug pattern responders. Relapse frequencies did not differ significantly between the two pattern groups, and survival estimates of time to relapse were virtually equal. This null result parallels a similar finding reported by Nierenberg et al. (2004), who also failed to find the hypothesized pattern differences for those continuing on mirtazepine in their adult study. By way of explanation, Nierenberg and his colleagues noted that "response patterns are inexact" (p. 1015), suggesting that both true drug and placebo response patterns are heterogeneous. The issue here is the accuracy of pattern analysis.

Accuracy of pattern analysis. Evidence suggests that the accuracy of pattern analysis is not symmetrical regarding true drug and placebo pattern designations. Data derived from the early Quitkin studies (1984, 1987) showed that in adults, the 'true drug' combination of *delay and persistence* correctly

identified patients on active drug (specificity = .81), but also yielded a number of false negatives (sensitivity = .50). This suggests that true drug pattern responding identifies a relatively homogeneous group who are uniformly drug responsive, even if some of their number are missing (misclassified as placebo-type responders). As such, membership in this group provides a useful prognostic indicator regarding relapse on fluoxetine or placebo. By contrast, placebo pattern responding comprises a more heterogeneous group, including many exhibiting unstable placebo effects (and thus subject to relapse with the loss of placebo effects), but also including spontaneous remitters as well as some drug responsive patients misclassified by the Quitkin pattern protocol; outcomes for this group are less predictable. Thus, by way of explaining the nearly equivalent relapse rates for true drug and placebo pattern responders continuing on fluoxetine, placebo pattern responders may do relatively well (better than hypothesized) because, as a more heterogeneous group, they have at least two mechanisms for 'staying well': (a) sustained spontaneous remission, or (b) 'missed' drug responsiveness that the marginally sensitive delay-and-persistence test failed to pick-up.

Clinical application of pattern analysis. An implicit assumption of pattern analysis is that patient improvement while on antidepressants is attributable to either drug responsiveness *or* to nonspecific mechanisms (e.g., placebo effects or spontaneous remission), but not both. The more likely

scenario, however, is that the course of depression is multidetermined, depending partly on the degree of drug response, but also on placebo effects, natural disease course, life events, and a host of other factors (Quitkin et al., 1993b). This raises the question of whether or not pattern analysis draws a distinction that is meaningful enough to be useful. The answer, probably yes, lies in the robust showing of an enhanced drug-placebo effect for true drug responders, tempered somewhat by the low sensitivity of the delay-and-persistence test, which misses a fair number of fluoxetine responders. Nonetheless, this does not detract from the useful clinical reality that for the largely homogeneous group of young patients who demonstrate delayed and persistent improvement, it is reasonably certain that fluoxetine will greatly reduce the risk of relapse. For the more heterogeneous group of placebo pattern responders, one can only say that fluoxetine provides no marked benefit *on average,* but will likely reduce the risk for relapse in some individuals (i.e., the undetected drug responsive patients).

Pattern analysis in drug trials. In the research arena, the prevalence of placebo responding in pediatric patients often complicates the task of demonstrating a significant treatment effect. The findings of this study suggest that in antidepressant efficacy studies where treatment effect is elusive, rerunning the data analysis using only true drug pattern responders will increase the chances of finding a treatment effect if one exists (i.e., the drug works), which may be useful in deciding whether to engage in further drug trials.

Modeling Relapse Prediction

In addition to investigating the validity of pattern analysis in pediatric depression, the present study also sought to examine response pattern and continuation treatment alongside other variables in the larger framework of relapse prediction. This not only allowed for controlling other relevant covariates to better isolate the effects of pattern and treatment, but also provided the means to identify which of the added covariates might also contribute to relapse.

Treatment. Following the definitive continuation study results of Emslie and his colleagues (2008; based on the same participants as the present study), continuation treatment was expected and found to be a significant predictor of risk for relapse. What is noteworthy in the present study, however, is the apparent strength of treatment as a predictor that emerged after controlling for other key variables, representing a near doubling of the hazard ratio from 2.10 to 4.02 (although the 95% confidence intervals overlap to a degree). For these study patients, fluoxetine clearly reduced the risk for relapse. This leaves virtually no doubt as to the efficacy of fluoxetine continuation in reducing the risk of relapse.

Gender. Unlike treatment, the emergence of gender as a strong predictor of relapse was unexpected (hazard ratio = 2.30; Wald $\chi^2 = 2.30$, $df = 1$, $p = .006$). It is well known that Cox regression models tend to fit the "training data" (the data from which they are generated) better than they fit external data (Harrell,

1996), and thus for unexpected findings, the magnitude of the parameters and *p*-values are usually viewed with caution until they can be independently verified. In this case, however, a recent study of relapse predictors among adults (McGrath et al., 2006) found a nearly identical and equally unexpected finding regarding gender, where gender even outperformed continuation treatment in predicting risk for relapse (females at greater risk). These strong findings for gender call for independent testing for both populations in near-term studies.

Pattern. Out of 10 terms in the full fixed Cox proportional hazards regression analysis, the pattern X treatment interaction term was the third strongest predictor, although it fell just short of significance (Wald $\chi^2 = 3.28$, $df = 1$, $p = .07$). While this leaves open the questions of whether pattern analysis provides sufficient incremental validity to be useful in future models of relapse prediction, the showing here is strong enough to warrant further study.

Other covariates. A number of other, exploratory covariates failed to show significance as predictors of risk for relapse.

The null findings with regard to family history of mood disorder as predictor of relapse were not surprising given the ubiquity of depression among first order relatives of the patients, with nearly three quarters of patients (76/102 = 74.5%) reporting positive family histories. What might be of more

interest in future studies is the question of whether fluoxetine response patterns run in families, which would support drug responsiveness as heritable.

Reverse neurovegetative symptoms, common in atypical depression, were found in about approximately 1 in 10 study patients (10/102 = 9.8%). Although the results were in the predicted direction—that is, these patients did equally poorly on fluoxetine and placebo (4/5 = 80% relapse for both treatments)—the numbers were too small for meaningful regression analysis. The question of whether or not reverse neurovegetative symptoms are associated with low fluoxetine efficacy in the prevention of relapse remains open.

While adult studies have shown chronicity to be a robust predictor of risk for relapse (McGrath et al., 2000, 2006), chronicity (represented here as episode duration) fell short of significance. The hazard ratios for episode duration in the full fixed and trimmed Cox regression models carried *p*-values of .13 and .06, respectively. This weaker showing may reflect the fact that young patients have had less historical opportunity for depressive symptoms to reach chronic status. This is supported by the observation that only 2 of 102 patients were eligible for the DSM chronic mood specifier, which requires that criteria for a major depressive episode be met continuously for at least 2 years. With chronic depression redefined as episode duration $\geq$ 48 weeks, 14 of 102 patients (13.7%) met the criterion, but chi square analysis showed no significant association with

frequency of relapse ($\chi^2 = 0.47$, $df = 1$, $p = .50$). It appears that in the pediatric population, chronicity does not strongly contribute to modeling risk for relapse.

In summary, Cox regression analysis confirmed continuation treatment (fluoxetine vs. placebo) as a strong predictor of relapse. Specifically, when isolating the effects of treatment by including (and thereby controlling for) other relevant covariates, the observed treatment effect was more pronounced. Gender showed unexpected strength in relapse prediction, paralleling a similar strong finding for gender in a large study of adults (McGrath et al., 2006). While the regression analysis showed the interaction of pattern X treatment to fall just short of significance, this finding warrants continued inclusion of pattern effects in pediatric studies of relapse prediction.

When the Cox model was rerun using the stricter definition of relapse ("full relapse," i.e., CDRS-R ≥ 40 only), the treatment and gender findings were nearly identical, and severity at randomization (CDRS-R score at week 12) also emerged as a strong predictor. Pattern effects, however, ceased to be relevant, likely due to the increase in truncated survival data. The CDRR study protocol called for the withdrawal of deteriorating patients at risk for relapse. Their removal prior to full relapse meant that they were never considered in the data analysis as anything but 'well', thus obscuring pattern effects. Still, a question

remains as to whether or not pattern analysis works better when the threshold of change (relapse) during continuation treatment is less stringent.

Other Findings

The findings regarding differential dropout rates were statistically significant, but because so many of these dropouts were triggered by clinical observations of worsening (rather than voluntary withdrawal), the results were not enlightening. So, while true drug responders exited the continuation phase more often and sooner than other patients, voluntary dropouts were too few to detect any discernible patterns. In future studies, it may be useful to look at the 'short patterns' of patients who drop out during the acute treatment phase.

Consistent with the hypothesis regarding abrupt improvement during acute treatment, all 19 such cases occurred in the first 2 weeks, designating them as placebo-type responders. However, improvement persisted in nearly all (18 of 19) of these cases, suggesting that they reflect spontaneous remission rather than actual placebo effects, which tend to less stable (Quitkin, Rabkin et al., 1991). Although not investigated here, it follows that those with early abrupt improvements may carry a better prognosis than other placebo pattern responders.

The finding that the performance of six-digit (8-week) response patterns is on a par with eight-digit (12-week) response patterns suggests that pharmacologic (true drug) responding is largely discernible by 8 weeks.

Limitations of Study

This was a post-hoc study for which acute pattern determination and all data analyses were conducted after key elements of outcome were already known (published in Emslie et al., 2008). A prospective study including a placebo arm in the acute phase would have provided a stronger design. However, as noted by Nierenberg et al. (2004), the open trial during the acute phase mirrors clinical practice and therefore provides a measure of ecological validity.

Perhaps the main shortcoming in the present study was starting with the assumption that adult response patterns are valid in pediatric patients, without benefit of prior research in young patients regarding typical patterns of response to active medication versus placebo (see Future Directions, below).

Future Directions

The present study suggests that pattern analysis can be usefully applied research and clinical settings involving pediatric depression, and therefore further research is warranted. Two possible lines of research are proposed here:

Response pattern validation. Early pattern analysis studies (Quitkin et al., 1984, 1987) started with placebo-controlled drug trials to establish how drug response differs from response to placebo, and only later were these prototypical

patterns used to classify improvement on active medication as likely due to pharmacologic versus other, nonspecific mechanisms. The present study started with the assumption that the Quitkin adult patterns apply to children as well, then proceeded to examine how well they predicted differential treatment effects in a placebo-controlled continuation trial to study relapse prevention. While the results here showed that pattern analysis contributes to the prediction of continuation relapse—and this alone provides a measure of validation for the method—it would be useful to examine the data from earlier pediatric fluoxetine trials to verify patterns directly for children and adolescents.

Refining pattern analysis. Pattern analysis in its present form is inexact, and it appears likely a fair number of drug responsive patients are misclassified as placebo-type responders. Several adult studies have demonstrated that early signs of drug effects (i.e., within the first week or two of treatment inception) are detectable, and although they do not usually reach the level of clinically significant improvement, they often provide early indications of positive outcome (Khan et al., 1989; Katz et al., 1997, 2004, 2006; Stassen et al., 1998). The CGI-I rating scale that forms the basis of pattern analysis may be too blunt an instrument to adequately characterize the sequence of symptom remission associated with drug versus nonspecific (nonpharmacologic) responding. It would be useful to look at data gathered at frequent early intervals in acute drug trials to track symptom

remission for each of the 17 items on the CDRS-R in order to determine which symptoms abate first for those who improve on antidepressant medication versus those who improve on placebo. Such information could be used to enhance the accuracy of pattern classifications. There is some evidence that response profiles may vary among antidepressant medications (Katz et al., 1997, 2006).

Pattern analysis in its current form is simple and therefore intuitively appealing for clinical decision making regarding continuation treatment (i.e., delayed and persistent improvement = true drug responding). For remitted young patients who exhibit this pattern, continuation treatment can be described as clearly beneficial in reducing the risk for relapse. Any refinements of pattern analysis are likely to sacrifice this simplicity in trade for greater accuracy.

Summary

The present study demonstrates that the Quitkin method of pattern analysis in adults can be usefully extended to pediatric depression. The following are the main conclusions of this study:

1. Pattern analysis differentiates true drug benefits from placebo effects in children and adolescents as effectively as it does in adults. This is most clearly seen in the enhanced treatment effect of fluoxetine versus placebo in reducing the risk for relapse in true drug pattern responders as compared with the absence of any significant treatment effect for placebo pattern responders.
2. Among pediatric patients, placebo pattern responders are, on average, a year younger than placebo pattern responders, which is consistent with prior reports of high rates of placebo responding in young children (Bridge et al., 2007), and intuitively consistent with developmental differences between young children and adolescents.
3. The accuracy of pattern analysis in differentiating drug benefits from placebo effects is not symmetric. The true drug response pattern successfully identifies a homogenous group of drug responsive patients, but it also misses some due to only modest sensitivity. By contrast, placebo pattern responders represent a more heterogeneous group, variously reflecting placebo effects, spontaneous remission, as

well as drug responsive patients who have been misdesignated. The true drug pattern therefore has greater prognostic value than the placebo pattern. In short, since (nearly) all true drug pattern responders probably are, in fact, drug responsive, treating clinicians can be confident that for those showing a delayed and persistent pattern, continuation treatment will substantially reduce their risk for relapse.

4. Cox regression shows treatment and gender to be strong predictors of risk for elapse, and the interaction term of pattern X treatment makes a respectable showing as a predictor ($p = .07$). Severity at randomization is a strong predictor for full relapse.
5. Exploratory analyses to examine the predictive value of depression chronicity, family history, and neurovegetative symptom status did not yield meaningful results.
6. Future research should be directed at refining the methods of pattern analysis to improve pattern classification.

APPENDIX

Table A. *Pattern Data*

Seq No.	Study ID	CGI–Improvement Ratings (by week)								Pattern String	Improvement		Pattern
		1	2	3	4	6	8	10	12		Timing	Stability	
1	RR002	2	2	2	2	1	2	1	1	1111 1111	early	persistent	placebo type
2	RR004	3	1	1	1	1	1	1	1	0111 1111	early	persistent	placebo type
3	RR008	2	2	2	1	1	1	1	1	1111 1111	early	persistent	placebo type
4	RR009	3	2	2	1	1	1	1	1	0111 1111	early	persistent	placebo type
5	RR012	3	3	2	1	1	1	1	1	0011 1111	delayed	persistent	true drug
6	RR016	3	3	3	2	2	2	2	2	0001 1111	delayed	persistent	true drug
7	RR024	2	2	2	2	1	1	2	1	1111 1111	early	persistent	placebo type
8	RR030	3	2	1	1	1	1	1	2	0111 1111	early	persistent	placebo type
9	RR032	2	2	1	1	1	1	2	1	1111 1111	early	persistent	placebo type
10	RR034	3	3	2	1	1	1	1	1	0011 1111	delayed	persistent	true drug
11	RR038	3	3	2	2	2	2	2	1	0011 1111	delayed	persistent	true drug
12	RR040	2	1	1	1	1	1	1	1	1111 1111	early	persistent	placebo type
13	RR041	2	2	2	2	2	2	1	1	1111 1111	early	persistent	placebo type
14	RR043	4	3	3	2	2	3	2	2	0001 1011	delayed	nonpersistent	placebo type
15	RR044	4	3	3	3	3	3	2	2	0000 0011	delayed	persistent	true drug
16	RR047	3	3	3	2	2	1	1	1	0001 1111	delayed	persistent	true drug
17	RR050	4	2	2	4	4	4	2	2	0110 0011	early	nonpersistent	placebo type
18	RR053	4	2	2	2	2	1	1	1	0111 1111	early	persistent	placebo type
19	RR056	3	2	1	1	1	2	2	1	0111 1111	early	persistent	placebo type
20	RR060	2	1	1	1	1	1	1	1	1111 1111	early	persistent	placebo type
21	RR061	2	2	2	1	1	1	1	1	1111 1111	early	persistent	placebo type
22	RR062	2	1	1	2	1	1	1	1	1111 1111	early	persistent	placebo type
23	RR063	3	2	1	1	1	1	1	1	0111 1111	early	persistent	placebo type
24	RR066	3	1	1	3	1	1	1	1	0110 1111	early	nonpersistent	placebo type
25	RR069	4	3	3	3	3	2	3	2	0000 0101	delayed	nonpersistent	placebo type
26	RR071	2	2	1	1	1	1	1	1	1111 1111	early	persistent	placebo type
27	RR074	3	2	1	1	1	1	1	1	0111 1111	early	persistent	placebo type
28	RR078	3	2	2	1	1	1	1	1	0111 1111	early	persistent	placebo type
29	RR083	4	3	3	3	3	3	3	1	0000 0001	delayed	indeterminate	indeterminate
30	RR084	3	3	2	2	1	1	1	1	0011 1111	delayed	persistent	true drug
31	RR087	3	3	2	2	1	1	1	1	0011 1111	delayed	persistent	true drug
32	RR089	4	3	2	2	2	2	1	1	0011 1111	delayed	persistent	true drug
33	RR092	4	3	3	2	3	2	-	1	0001 0111	delayed	nonpersistent	placebo type
34	RR093	4	3	2	2	1	1	1	1	0011 1111	delayed	persistent	true drug
35	RR094	3	3	3	3	2	1	1	1	0000 1111	delayed	persistent	true drug
36	RR095	4	3	2	2	2	1	1	1	0011 1111	delayed	persistent	true drug
37	RR096	3	3	3	3	3	2	1	1	0000 0111	delayed	persistent	true drug
38	RR097	3	3	3	2	2	1	1	1	0001 1111	delayed	persistent	true drug
39	RR098	3	2	2	2	2	2	1	2	0111 1111	early	persistent	placebo type
40	RR102	3	3	3	3	2	2	2	2	0000 1111	delayed	persistent	true drug
41	RR105	3	3	2	2	2	1	1	1	0011 1111	delayed	persistent	true drug
42	RR106	4	2	2	2	2	1	1	1	0111 1111	early	persistent	placebo type
43	RR107	4	3	2	2	2	2	1	1	0011 1111	delayed	persistent	true drug
44	RR108	4	3	3	2	2	2	1	1	0001 1111	delayed	persistent	true drug
45	RR110	4	3	3	2	1	1	1	1	0001 1111	delayed	persistent	true drug
46	RR113	3	2	1	1	1	1	1	1	0111 1111	early	persistent	placebo type
47	RR128	4	3	2	2	2	2	1	1	0011 1111	delayed	persistent	true drug
48	RR130	3	2	3	3	3	2	1	1	0100 0111	early	nonpersistent	placebo type
49	RR131	4	3	3	2	2	1	1	1	0001 1111	delayed	persistent	true drug
50	RR140	4	3	3	3	3	3	2	2	0000 0011	delayed	persistent	true drug
51	RR149	4	3	3	2	2	2	2	1	0001 1111	delayed	persistent	true drug

Table A. *Pattern Data (cont'd.)*

Seq No.	Study ID	CGI–Improvement Ratings (by week)								Pattern String	Improvement		Pattern
		1	2	3	4	6	8	10	12		Timing	Stability	
52	RR155	3	2	2	2	2	2	1	1	0111 1111	early	persistent	placebo type
53	RR159	4	3	2	2	1	1	1	1	0011 1111	delayed	persistent	true drug
54	RR165	3	3	3	2	1	1	1	1	0001 1111	delayed	persistent	true drug
55	RR171	3	3	3	3	2	1	1	1	0000 1111	delayed	persistent	true drug
56	RR172	3	3	2	2	1	1	1	1	0011 1111	delayed	persistent	true drug
57	RR175	4	4	3	3	2	2	2	2	0000 1111	delayed	persistent	true drug
58	RR179	4	3	3	3	1	2	1	1	0000 1111	delayed	persistent	true drug
59	RR183	3	2	3	2	2	1	1	1	0101 1111	early	nonpersistent	placebo type
60	RR189	4	4	4	3	3	2	2	2	0000 0111	delayed	persistent	true drug
61	RR190	3	3	2	2	1	1	1	1	0011 1111	delayed	persistent	true drug
62	RR191	3	3	3	2	3	3	3	2	0001 0001	delayed	nonpersistent	placebo type
63	RR192	4	3	2	1	1	1	1	1	0011 1111	delayed	persistent	true drug
64	RR194	2	1	1	1	1	-	2	2	1111 1111	early	persistent	placebo type
65	RR196	3	2	2	2	2	2	1	1	0111 1111	early	persistent	placebo type
66	RR197	4	3	3	2	2	2	1	1	0001 1111	delayed	persistent	true drug
67	RR203	4	2	2	2	1	1	-	1	0111 1111	early	persistent	placebo type
68	RR204	3	2	3	2	1	-	1	1	0101 1111	early	nonpersistent	placebo type
69	RR206	3	2	2	1	1	1	2	2	0111 1111	early	persistent	placebo type
70	RR215	4	3	3	3	3	2	2	2	0000 0111	delayed	persistent	true drug
71	RR220	3	2	1	1	1	1	2	1	0111 1111	early	persistent	placebo type
72	RR223	3	2	2	2	2	2	2	2	0111 1111	early	persistent	placebo type
73	RR224	4	3	3	4	2	2	1	1	0000 1111	delayed	persistent	true drug
74	RR225	4	4	3	3	3	3	2	2	0000 0011	delayed	persistent	true drug
75	RR228	4	3	2	1	2	2	2	2	0011 1111	delayed	persistent	true drug
76	RR230	3	2	-	2	2	2	1	1	0111 1111	early	persistent	placebo type
77	RR237	3	2	2	2	1	1	1	1	0111 1111	early	persistent	placebo type
78	RR240	3	2	2	1	1	1	1	1	0111 1111	early	persistent	placebo type
79	RR242	3	2	2	1	1	2	2	1	0111 1111	early	persistent	placebo type
80	RR243	3	-	3	3	2	2	1	1	0000 1111	delayed	persistent	true drug
81	RR245	4	3	3	2	2	2	2	2	0001 1111	delayed	persistent	true drug
82	RR246	3	2	2	1	1	1	1	2	0111 1111	early	persistent	placebo type
83	RR247	3	2	-	2	1	1	1	1	0111 1111	early	persistent	placebo type
84	RR250	3	3	2	1	1	1	1	1	0011 1111	delayed	persistent	true drug
85	RR251	3	2	1	1	1	1	1	1	0111 1111	early	persistent	placebo type
86	RR256	4	3	2	2	2	2	2	2	0011 1111	delayed	persistent	true drug
87	RR263	3	2	1	1	1	1	1	1	0111 1111	early	persistent	placebo type
88	RR265	3	2	2	2	1	1	1	1	0111 1111	early	persistent	placebo type
89	RR266	2	2	1	1	1	1	1	1	1111 1111	early	persistent	placebo type
90	RR269	4	3	3	3	3	3	2	2	0000 0011	delayed	persistent	true drug
91	RR270	3	3	2	2	2	1	1	1	0011 1111	delayed	persistent	true drug
92	RR271	4	2	2	1	1	1	1	1	0111 1111	early	persistent	placebo type
93	RR277	4	4	3	2	1	1	1	1	0001 1111	delayed	persistent	true drug
94	RR287	4	2	2	1	1	1	1	1	0111 1111	early	persistent	placebo type
95	RR289	3	2	2	2	2	2	2	2	0111 1111	early	persistent	placebo type
96	RR302	3	3	3	2	2	2	2	1	0001 1111	delayed	persistent	true drug
97	RR304	2	2	1	2	1	1	-	1	1111 1111	early	persistent	placebo type
98	RR305	4	-	3	3	2	2	2	1	0000 1111	delayed	persistent	true drug
99	RR307	4	3	3	2	2	2	2	2	0001 1111	delayed	persistent	true drug
100	RR308	3	3	2	2	2	2	2	1	0011 1111	delayed	persistent	true drug
101	RR325	4	4	3	3	2	2	2	2	0000 1111	delayed	persistent	true drug
102	RR331	4	4	-	3	3	2	2	2	0000 0111	delayed	persistent	true drug

REFERENCES

American Psychiatric Association. (1987). *Diagnostic and statistical manual of mental disorders (3rd ed., rev.)*. Washington, D.C.: American Psychiatric Association.

American Psychiatric Association. (2000). *Diagnostic and statistical manual of Mental disorders (4th ed., text revision)* (4th, text ed.). Washington, D.C.: American Psychiatric Association.

Andreasen, N. C., Rice, J., Endicott, J., Reich, T., & Coryell, W. (1986). The family history approach to diagnosis: How useful is it? *Archives of General Psychiatry, 43*(5), 421-429.

Birmaher, B., Ryan, N. D., Williamson, D. E., Brent, D. A., & Kaufman, J. (1996b). Childhood and adolescent depression: A review of the past 10 years. Part II. *Journal of the American Academy of Child and Adolescent Psychiatry, 35*(12), 1575-1583.

Birmaher, B., Ryan, N. D., Williamson, D. E., Brent, D. A., Kaufman, J., Dohl, R. E., et al. (1996a). Childhood and adolescent depression: A review of the past 10 years. Part I. *Journal of the American Academy of Child and Adolescent Psychiatry, 35*(11), 1427-1439.

Brent, D. A., Holder, D., Kolko, D., Birmaher, B., Baugher, M., Roth, C., et al. (1997). A clinical psychotherapy trial for adolescent depression comparing

cognitive, family, and supportive therapy. *Archives of General Psychiatry, 54*(9), 877-885.

Bridge, J. A., Iyengar, S., Salary, C. B., Barbe, R. P., Birmaher, B., Pincus, H. A., et al. (2007). Clinical response and risk for reported suicidal ideation and suicide attempts in pediatric antidepressant treatment: A meta-analysis of randomized controlled trials. *Journal of the American Medical Association, 297*(15), 1683-1696.

Brooks, S. J., & Kutcher, S. (2001). Diagnosis and measurement of adolescent depression: A review of commonly utilized instruments. *Journal of Child and Adolescent Psychopharmacology, 11*(4), 341-376.

Byrne, S. E., & Rothschild, A. J. (1998). Loss of antidepressant efficacy during maintenance therapy: Possible mechanisms and treatments. *Journal of Clinical Psychiatry, 59*(6), 279-288.

Cheung, A. H., Emslie, G. J., & Mayes, T. L. (2006). The use of antidepressants to reat depression in children and adolescents. *Canadian Medical Association Journal, 174*(2), 193-200.

Cohen, B. M., & Baldessarini, R. J. (1985). Tolerance to therapeutic effects of antidepressants. *American Journal of Psychiatry, 142*, 489-490.

Cohen, J. (1988). *Statistical Power Analysis for the Behavioral Sciences* (2nd ed.). Hillsdale, NJ: Lawrence Erlbaum Associates.

Concato, J., Peduzzi, P., Holford, T. R., & Feinstein, A. R. (1995). Importance of events per independent variable in proportional hazards analysis: I. Background, goals, and general strategy. *Journal of Clinical Epidemiology, 48*(12), 1495-1501.

Costello, E. J., Erkanli, A., & Angold, A. (2006). Is there an epidemic of child or adolescent depression? *Journal of Child Psychology and Psychiatry, 47*(12), 1263-1271.

Cox, D. R. (1972). Regression models and life tables (with discussion). *Journal of the Royal Statistical Society, B34*, 187-220.

Derivan, A., & Entsuah, A. R. (1995). Venlafaxine: Measuring the onset of antidepressant action. *Psychopharmacology Bulletin, 34*, 437-445.

Derksen, S., & Keselman, H. J. (1992). Backward, forward and stepwise automated subset selection algorithms: Frequency of obtaining authentic and noise variables. *British Journal of Mathematical and Statistical Psychology, 45*, 265-282.

Dunlop, S. R., Dornseif, B. E., Wernicke, J. F., & Potvin, J. H. (1990). Pattern analysis shows beneficial effect of fluoxetine treatment in mild depression. *Psychopharmacology Bulletin, 26*(2), 173-180.

Dunn, V., & Goodyer, I. M. (2006). Longitudinal investigation into childhood- and adolescence-onset depression: Psychiatric outcome in early adulthood. *British Journal of Psychiatry, 188*(3), 216-222.

Dworkin, R. H., Katz, J., & Gitlin, M. J. (2005). Placebo response in clinical trials of depression and its implications for research on chronic neuropathic pain. *Neurology (Suppl. 4), 65*(12), S7-S19.

Emslie, G. J., Findling, R. L., Yeung, P. P., Kunz, N. R., & Yunfend, L. (2007). Venlafaxine ER for the treatment of pediatric subjects with depression: Results of two placebo-controlled trials. *Journal of the American Academy of Child and Adolescent Psychiatry, 46*(4), 479-488.

Emslie, G. J., Heiligenstein, J. H., Hoog, S. L., Wagner, K. D., Findling, R. L., McCracken, J. T., et al. (2004). Fluoxetine treatment for prevention of relapse of depression in children and adolescents: A double-blind, placebo-controlled study. *Journal of the American Academy of Child and Adolescent Psychiatry, 43*(11), 1397-1405.

Emslie, G. J., Heiligenstein, J. H., Wagner, K. D., Hoog, S. L., Ernest, D. E., Brown, E., et al. (2002). Fluoxetine for acute treatment of depression in children and adolescents: A placebo-controlled, randomized clinical trial. *Journal of the American Academy of Child and Adolescent Psychiatry, 41*(10), 1205-1215.

Emslie, G. J., Kennard, B. D., Mayes, T. L., Nightingale-Teresi, J., Carmody, T., Hughes, C. W., et al. (2008). Fluoxetine vs. placebo to prevent relapse in MDD in children and adolescents. *American Journal of Psychiatry, 165*.

Emslie, G. J., Mayes, T. L., & Ruberu, M. (2005). Continuation and maintenance therapy of early-onset major depressive disorder. *Pediatric Drugs, 7*(4), 203-217.

Emslie, G. J., Rush, A. J., Weinberg, W. A., Kowatch, R. A., Carmody, T., & Mayes, T. L. (1998). Fluoxetine in child and adolescent depression: Acute and maintenance treatment. *Depression and Anxiety, 7*(1), 32-39.

Emslie, G. J., Rush, A. J., Weinberg, W. A., Kowatch, R. A., Hughes, C., Carmody, T., et al. (1997). A double-blind, randomized, placebo-controlled trial of fluoxetine in children and adolescents with depression. *Archives of General Psychiatry, 54*, 1031-1037.

Fieve, R. R., Goodnick, P. J., Peselow, E. D., Barouche, F., & Schlegel, A. (1986). Pattern analysis of antidepressant response to fluoxetine. *Journal of Clinical Psychiatry, 47*, 560-562.

Frank, E., Prien, R. F., Jarrett, R. B., Keller, M. B., Kupfer, D. J., Lavori, P. W., et al. (1991). Conceptualization and rationale for consensus definitions of terms in major depressive disorder: Remission, recovery, relapse, and recurrence. *Archives of General Psychiatry, 48*(9), 851-855.

Garson, G. D. (n.d.). Cox regression [Electronic Version]. *Statnotes: Topics in Multivariate Analysis.* Retrieved April 23, 2008, from http://www2.chass.ncsu.edu/garson/pa765.cox.htm.

Geddes, J. R., Carney, S. M., Davies, C., Furukawa, T., Kupfer, D. J., Frank, E., et al. (2003). Relapse prevention with antidepressant drug treatment in depressive disorders: A systematic review. *Lancet, 361*, 653-661.

Gibbons, R. D., Hur, K., Bhaumik, D. K., & Mann, J. J. (2005). The relationship between antidepressant medication use and rate of suicide. *Archives of General Psychiatry, 62*(2), 165-172.

Guy, W. (1976). Clinical Global Impressions scale. In *ECDEU assessment manual for psychopharmacology, revised* (pp. 218-222). Rockville, MD: National Institute of Mental Health.

Hamilton, M. (1960). A rating scale for depression. *Journal of neurology, neurosurgery and psychiatry, 23*, 56-62.

Hammad, T. A., Laughren, T., & Racoosin, J. (2006). Suicidality in pediatric patients treated with antidepressant drugs. *Archives of General Psychiatry, 63*(3), 332-339.

Hammerness, P. G., Vivas, F. M., & Geller, D. A. (2006). Selective serotonin reuptake inhibitors in pediatric psychopharmacology: A review of the evidence. *Journal of Pediatrics, 148*(2), 158-165.

Harrell, F. E., Jr. (1996). Multivariable prognostic models: Issues in developing models, evaluating assumptions and adequacy, and measuring and reducing errors. *Statistics in Medicine, 15*(4), 361-387.

Harrell, F. E., Jr. (2001). *Regression modeling strategies: With applications to linear models, logistic regression, and survival analysis* (1st edition ed.). New York: Springer.

Harrell, F. E., Jr. (2002, August 1). Re: Stepwise and significance: Message posted to http://lists.biostat.wustl.edu.

Harrell, F. E., Jr., Lee, K. L., Califf, R. M., Pryor, D. B., & Rosati, R. A. (1984). Regression modelling strategies for improved prognostic prediction. *Statistics in Medicine, 3*(2), 143-152.

Harrington, R., Whittaker, J., Shoebridge, P., & Campbell, F. (1998). Systematic review of efficacy of cognitive behaviour therapies in childhood and adolescent depressive disorder. *British Medical Journal, 316*, 1559-1563.

Hazell, P., O'Connell, C., Heathcote, D., Robertson, J., & Henry, D. (1995). Efficacy of tricyclic drugs in treating child and adolescent depression: A meta-analysis. *British Medical Journal, 310*, 897-901.

Hughes, C. W., Emslie, G., Kowatch, R., Weinberg, W., Rintelmann, J., & Rush, A. J. (2000). Clinican, Parent, and Child Prediction of Medication or Placebo in Double-Blind Depression Study. *Neuropsychopharmacology, 23*(5), 591-594.

Hughes, C. W., Emslie, G. J., Crismon, M. L., Posner, K., Birmaher, B., Ryan, N., et al. (2007). Texas Children's Medication Algorithm Project: Update from Texas Consensus Conference Panel on Medication Treatment of

Childhood Major Depressive Disorder. *Journal of the American Academy of Child and Adolescent Psychiatry, 46*(6), 667-686.

Jones, M. B., & Ainslie, J. D. (1966). Value of a placebo wash-out. *Diseases of the Nervous System, 27*, 393-396.

Journal of Clinical Psychiatry. (2001). Early onset of antidepressant action [Special issue]. *JCP, 62*(Suppl. 4).

Kane, E. P., Fagan, H. B., & Wolf, D. G. (2007). Sould we use SSRIs to treat adolescents with depression? *Journal of Family Practice, 56*(9), 759-760.

Katz, M. M., Bowden, C. L., Berman, N., & Frazer, A. (2006). Resolving the onset of antidepressants' clinical actions: Critical for clinical practice and new drug development. *Journal of Clinical Psychopharmacology, 26*(6).

Katz, M. M., Koslow, S. H., & Frazer, A. (1997). Onset of antidepressant activity: Reexamining the structure of depression and multiple actions of drugs. *Depression and Anxiety, 4*(6), 257-267.

Katz, M. M., Tekell, J. L., Bowden, C. L., Brannan, S., Houston, J. P., Berman, N., et al. (2004). Onset and early behavioral effects of pharmacologically different antidepressants and placebo in depression. *Neuropsychopharmacology, 29*(3), 566-579.

Kaufman, J., Birmaher, B., Brent, D., Rao, U., Flynn, C., Moreci, P., et al. (1997). Schedule for Affective Disorders and Schizophrenia for School-Age Children—Present and Lifetime Version (K-SADS-PL): Initial reliability

and validity data. *Journal of the American Academy of Child and Adolescent Psychiatry, 36*(7), 980-988.

Kaufman, J., Birmaher, B., Brent, D., Rao, U., & Ryan, N. (1996). *Schedule for Affective Disorders and Schizophrenia for School-Age Children: Present and Lifetime version (K-SADS-PL)*. Pittsburgh: Western Psychiatric Institute and Clinic.

Kennard, B., Silva, S., Vitiello, B., Curry, J., Kratochvil, C., Simons, A., et al. (2006). Remission and residual symptoms after short-term treatment in the Treatment of Adolescents With Depression Study (TADS). *Journal of the American Academy of Child and Adolescent Psychiatry, 45*(12), 1404-1411.

Khan, A., & Brown, W. A. (2001). The placebo enigma in antidepressant clinical trials. *Journal of Clinical Psychopharmacology, 21*(2), 123-125.

Khan, A., Cohen, S., Dager, t., Avery, D. H., & Dunner, D. L. (1989). Onset of response in relation to outcome in depressed outpatients with placebo and imipramine. *Journal of Affective Disorders, 17*(1), 33-38.

Klein, D. F., Gittelman, R., Quitkin, F., & Rifkin, A. (1980). *Diagnosis and drug treatment of psychiatric disorders: Adults and Children* (2nd ed.). Baltimore: Williams & Wilkins.

Klerman, G. I., & Cole, J. O. (1965). Clinical pharmacology of imipramine and related antidepressants. *Psychopharmacology Review, 17*, 101-141.

Kowatch, R. A., Carmody, T. J., Emslie, G. J., Rintelmann, J. W., Hughes, C. W., & Rush, A. J. (1999). Prediction of response to fluoxetine and placebo in children and adolescents with major depression: A hypothesis generating study. *Journal of Affective Disorders, 54*((3)).

Kuhn, R. (1958). The treatment of depressive states with G 22355 (imipramine hydrochloride). *American Journal of Psychiatry, 115*, 459-464.

Kutcher, S. P. (1999). Pharmacoptherapy of depression: A review of current evidence and practical clinical direction. In C. A. Essau & F. Petermann (Eds.), *Depressive disorders in children and adolescents: Epidemiology, risk factors, and treatment* (pp. 437-458). Northvale, NJ: Jason Aronson.

Lefkowitz, M. M., & Burton, N. (1978). Childhood depression: A critique of the concept. *Psychological Bulletin, 85*(4), 716-726.

Mann, J. J., Emslie, G., Baldessarini, R. J., Beardslee, W., Fawcett, J. A., Goodwin, F. K., et al. (2006). ACNP Task Force report on SSRIs and suicidal behavior. *Neuropsychopharmacology, 31*(3), 473-492.

Mayes, T. L., Tao, R., Rintelmann, J. W., Carmody, T., Hughes, Carroll W., Kennard, B. D., Stewart, S. M., et al. (2007). Do children and adolescents have differential response rates in placebo-controlled trials of fluoxetine? *CNS Spectrums, 12*(2), 147-154.

McGrath, P. J., Stewart, J. W., Petkova, E., Quitkin, F. M., Amsterdam, J. D., Fawcett, J., et al. (2000). Predictors of relapse during fluoxetine

continuation or maintenance treatment of major depression. *Journal of Clinical Psychiatry, 61*(7), 518-524.

McGrath, P. J., Stewart, J. W., Quitkin, F. M., Chen, Y., Alpert, J. E., Fava, M., et al. (2006). Predictors of relapse in a prospective study of fluoxetine treatment of major depression. *American Journal of Psychiatry, 163*(9), 1542-1546.

Meadows, M. (2003). Drug research and children. *FDA Consumer, 37*.

Montgomery, S. A. (1994). Long-term treatment of depression. *British Journal of Psychiatry Supplement, 26S*.

Montgomery, S. A., Dufour, H., Brion, S., Gailledreau, J., Laqueille, X., Ferry, G., et al. (1988). The prophylactic efficacy of fluoxetine in unipolar depression. *British Journal of Psychiatry, 153*(suppl.), 69-76.

Moreno, C., Arango, C., Parellada, M., Shaffer, D., & Bird, H. (2007). Antidepressants in child and adolescent depression: Where are the bugs? *Acta Psychiatrica Scandinavica, 115*, 194-195.

Mrazek, D. A., & Masterson, J. (1985). Family Global Assessment Scale (FGAS): Initial reliability and validity. *Scientific Proceedings for the 32nd Annual Meeting of the American Academy of Child and Adolescent psychiatry (San Antonio, TX, October 23-27)*, p. 38.

National Institute of Mental Health. (1985). Clinical Global Impressions Scale. *Psychopharmacology Bulletin, 21*(4), 839-843.

National Institute of Mental Health. (2007). Antidepressant medication for children and adolescents: Information for parents and caregivers [Electronic Version]. Retrieved September 8, 2007, from http://www.nimh.nih.gov/healthinformation/antidepressant_child.cfm.

Nierenberg, A. A., Farabaugh, A. H., Alpert, J. E., Gordon, J., Worthington, J. J., Rosenbaum, J. F., et al. (2000). Timing of onset of antidepressant response with fluoxetine treatment. *American Journal of Psychiatry, 157*(9), 1423-1428.

Nierenberg, A. A., Quitkin, F. M., Kremer, C., Keller, M. B., & Thase, M. E. (2004). Placebo-controlled continuation treatment with mirtazapine: Acute pattern of response predicts relapse. *Neuropsychopharmacology, 29*, 1012-1018.

Olfson, M., Shaffer, D., Marcus, S. C., & Greenberg, T. (2003). Relationship between antidepressant medication treatment and suicide in adolescents. *Archives of General Psychiatry, 60*(10), 978-982.

Pavuluri, M., & Birmaher, B. (2004). A practical guide to using ratings of depression and anxiety in child psychiatric practice. *Current Psychiatry Reports, 6*(2), 108-116.

Peduzzi, P., Concato, J., Feinstein, A. R., & Holford, T. R. (1995). Importance of events per independent variable in proportional hazards regression

analysis: II. Accuracy and precision of regression estimates. *Journal of Clinical Epidemiology, 48*(12), 1503-1510.

Pollack, B. (1959). Clinical findings in the use of Tofrānil in depressive and other psychiatric states. *Journal of the American Psychiatric Association, 116*(4).

Posternak, M. A., & Zimmerman, M. (2005). Is there a delay in the antidepressant effect? *Journal of Clinical Psychiatry, 66*(2), 148-158.

Poznanski, E. O., Grossman, J. A., Buchsbaum, Y., Banegas, M., Freeman, L., & Gibbons, R. (1984). Preliminary studies of the reliability and validity of the Children's Depression Rating Scale. *Journal of the American Academy of Child and Adolescent Psychiatry, 23*(2), 191-197.

Poznanski, E. O., & Mokros, H. B. (1996). *Children's Depression Rating Scale, Revised (CDRS-R): Manual*. Los Angeles: Western Psychological Services.

Prien, R. F., Carpenter, L. L., & Kupfer, D. J. (1991). The definition and operational criteria for treatment outcome of major depressive disorder: A review of the current research literature. *Archives of General Psychiatry, 48*(9), 796-800.

Priest, R. G., Hawley, C. J., Kibel, D., Kurian, T., Montgomery, S. A., Patel, A. G., et al. (1996). Recovery from depressive illness does fit an exponential model. *Journal of Clinical psychopharmacology, 16*(6), 420-424.

Quitkin, F. M., McGrath, P. J., Rabkin, J. G., Stewart, J. W., Harrison, W., Ross, D. C., et al. (1991). Different types of placebo response in patients receiving antidepressants. *American Journal of Psychiatry, 148*(2), 197-203.

Quitkin, F. M., Rabkin, J. D., Markowitz, J. M., Stewart, J. W., McGrath, P. J., & Harrison, W. (1987). Use of pattern analysis to identify true drug response: A replication. *Archives of General Psychiatry, 44*, 259-264.

Quitkin, F. M., & Rabkin, J. G. (1981). Methodological problems in studies of depressive disorder: Utility of the discontinuation design. *Journal of Clinical Psychopharmacology, 1*(5), 283-288.

Quitkin, F. M., Rabkin, J. G., Ross, D., & Stewart, J. W. (1984). Identification of true drug response to antidepressants: Use of pattern analysis. *Archives of General Psychiatry, 41*, 782-786.

Quitkin, F. M., Rabkin, J. G., Stewart, J. W., McGrath, P. J., Harrison, W., Ross, D. C., et al. (1991). Heterogeneity of clinical response during placebo treatment. *American Journal of Psychiatry, 148*(2), 193-196.

Quitkin, F. M., Stewart, J. W., McGrath, P. J., Nunes, E., Ocepek-Welikson, K., Tricamo, E., et al. (1993a). Further evidence that a placebo response to antidepressants can be identified. *American Journal of Psychiatry, 150*(4), 566-570.

Quitkin, F. M., Stewart, J. W., McGrath, P. J., Nunes, E., Ocepek-Welikson, K., Tricamo, E., et al. (1993b). Loss of drug effects during continuation therapy. *American Journal of Psychiatry, 150*(4), 562-565.

Rabkin, J. G., McGrath, P. J., Quitkin, F. M., Tricamo, E., Stewart, J. W., & Klein, D. F. (1990). Effects of pill-giving on maintenance of placebo response in patients with chronic mild depression. *American Journal of Psychiatry, 147*(12), 1622-1626.

Rickels, K., Derivan, A., Entsuah, R., Miska, S., & Rudolph, R. (1995). Rapid onset antidepressant activity with venlafaxine treatment. *Depression, 146*, 146-153.

Rothschild, R., & Quitkin, F. M. (1992). Review of the use of pattern analysis to differentiate true drug and placebo responses. *Psychotherapy and Psychosomatics, 58*, 170-177.

Rush, A. J., First, M. B., & Blacker, D. (Eds.). (2008). *Handbook of psychiatric measures* (2nd ed.). Washington, D.C.: American Psychiatric Association.

Schneider, U., Borsutzky, M., Seifert, J., Leweke, F. M., Huber, T. J., Tollnik, J. D., et al. (2002). Reduced binocular depth inversion in schizophrenic patients. *Schizophrenia Research, 53*(1-2), 101-108.

Schulterbrandt, J. G., & Raskin, A. (Eds.). (1977). *Depression in children: Diagnosis, treatment, and conceptual models*. New York: Raven Press.

Shaffer, D., Gould, M. S., Brasic, J., Ambrosini, P., Fisher, P., & Aluwahlia, S. (1983). A children's global assessment scale (CGAS). *Archives of General Psychiatry, 40*(11), 1228-1231.

Small, J. G., Milstein, V., Kellams, J. J., & Small, I. F. (1981). Comparative onset of improvement in depressive symptomatology with drug treatment, electroconvulsive therapy and placebo. *Journal of Clinical Psychopharmacology, 1*(suppl.), 62-69.

Spotswood, S. (2004). FDA committee advises strong antidepressant warning for kids [Electronic Version]. *U.S. Medicine*. Retrieved January 20, 2008, from http://www.usmedicine.com/article.cfm?articleID=954&issueID=67.

Stassen, H. H., & Angst, J. (1998). Delayed onset of action of antidepressants: Fact or fiction? *CNS Drugs, 9*(3), 177-184.

Stewart, J. W., Quitkin, F. M., McGrath, P. J., Amsterdam, J., Fava, M., Fawcett, J., et al. (1998). Use of pattern analysis to predict differential relapse of remitted patients with major depression during 1 year of treatment with fluoxetine or placebo. *Archives of General Psychiatry, 55*, 334-343.

TADS Team. (2007). The Treatment for Adolescents With Depression Study (TADS): Long-term effectiveness and safety outcomes. *Archives of General Psychiatry, 64*(10), 1132-1144.

Taylor, M. J., Freemantle, N., Geddes, J. R., & Bhagwagar, Z. (2006). Early onset of selective serotonin reuptake inhibitor antidepressant action. *Archives of General Psychiatry, 63*(11), 1217-1223.

Treatment for Adolescents With Depression Study [TADS] Team. (2004). Fluoxetine, cognitive-behavioral therapy, and their combination for adolescents with depression: Treatment for Adolescents With Depression Study (TADS) randomized controlled trial. *Journal of the American Medical Association, 292*(7), 807-820.

United States Food and Drug Administration. (2007). Questions and answers on antidepressant use in children, adolescents, and adults [Electronic Version]. Retrieved September 10, 2007 from http://www.fda.gov/cder/drug/antidepressants/QA20070502.htm#top.

Vitiello, B. (2007). Research in child and adolescent psychopharmacology: Recent accomplishments and new challenges. *Psychopharmacology, 191*(1), 5-13.

Vitiello, B., Calderoni, D., & Mazzone, L. (2006). Pharmacologic treatment of children and adolescents with major depressive disorder. In B. Vitiello, G. Masi & D. Marazziti (Eds.), *Handbook of child and adolescent psychopharmacology*. Abingdon, UK: Informa Healthcare.

Vittinghoff, E., & McColloch, C. E. (2007). Relaxing the rule of ten events per variable in logistic and Cox regression. *American Journal of Epidemiology, 165*(6), 710-718.

Wagner, K. D. (2005). Pharmacoptherapy for major depression in children and adolescents. *Progress in Neuro-Psychopharmacology & Biological Psychiatry, 29*(5), 819-826.

Wagner, K. D., Ambrosini, P., Rynn, M., Wohlber, C., Yang, R., Greenbaum, M. S., et al. (2003). Efficacy of sertraline in the treatment of children and adolescents with major depressive disorder. *Journal of the American Medical Association, 290*(8), 1033-1041.

Wagner, K. D., Robb, A. S., Findling, R. L., Jin, J., Gutierrez, M. M., & Heydorn, W. E. (2004). A randomized, placebo-controlled trial of citalopram for the treatment of major depressive disorder in children and adolescents. *American Journal of Psychiatry, 161*(6), 1079-1083.

Walsh, B. T., Seidman, S. N., Sysko, R., & Gould, M. (2002). Placebo response in studies of major depression. *Journal of the American Medical Association, 287*(14), 1840-1847.

Watanabe, N., Hunot, V., Omori, I. M., Churchill, R., & Furukawa, T. A. (2007). Psychotherapy for depression among children and adolescents: A systematic review. *Acta Psychiatrica Scandinavica, 116*(2).

Wechsler, D. (1991). Wechsler Intelligence Scale for Children—Third Edition. San Antonio, TX: The Psychological Corporation.

Weissman, M. M., & Shaffer, D. (1998). Introduction: Youthful depression special issue. *Depression and Anxiety, 7*, 1-2.

Weisz, J. R., McCarty, C. A., & Valeri, S. M. (2006). Effects of psychotherapy for depression in children and adolescents: A meta-analysis. *Psychological Bulletin, 132*(1), 132-149.

Weisz, J. R., Valeri, S. M., McCarty, C. A., & Moore, P. S. (1999). Interventions for depression: Features, effects and future directions. In C. A. Essau & F. Petermann (Eds.), *Depressive disorders in children and adolescents: Epidemiology, risk factors, and treatment* (pp. 383-435). Northvale, NJ: Jason Aronson.

Zimmerman, M., & Thongy, T. (2007). How often do SSRIs and other new-generation antidepressants lose their effect during continuation treatment? Evidence suggesting the rate of true tachyphylaxis during continuation treatment is low. *Journal of Clinical Psychiatry, 68*(8), 1271-1276.