Physician Academy Series
Educational Tools For Clinicians

Key words:
RCTs
antidepressants
drug efficacy
primary care
comparative
drug analysis

A Discussion on the Methodologies Used to Evaluate Antidepressants

MODERATOR:

Jan A. Fawcett, MD

DISCUSSANTS:

Norman Sussman, MD

Jeffrey H. Newcorn, MD

Also included in this program:

Lecture Slide Series

Free Interactive CD-ROM

Free Interactive CD-ROM Order Form

The contents of this Physician Academy Series including the **Lecture Slide Series** are available in an interactive format with PowerPoint® slides on an easy-to-use CD-ROM. The CD-ROM is compatible with both PC/Windows and Mac OS systems: no plug-ins are required. Just complete this form. It's that simple! This CD-ROM is offered to the medical community to encourage the rapid and proper dissemination of medical information.

The sponsor of this Physician Academy Series, Wyeth, has funded 1,000 CD-ROMs via an unrestricted educational grant for interested readers. For those first 1,000 requesters, a CD-ROM will be provided free of charge.

While the copyright of these slides lies with MedWorks Media LLC, the information on the slides, as well as the slides themselves, may be used by readers under the "fair use" doctrine with our blessing. Thus, a use for a proper purpose, such as teaching or lecturing, does not constitute infringement.

Please fill out the form below and fax it to 212.343.2225, or mail it to MedWorks Media, 375 West Broadway, Suite 501, New York, NY 10012. Please allow 3 weeks for delivery.

Name: _____

Address: _____

_____

_____

E-mail: _____

Specialty: _____

Date: _____

Thank you.

A Discussion on the Methodologies Used to Evaluate Antidepressants

Jan A. Fawcett, MD, is associate professor of psychiatry at Harvard Medical School in Boston, MA. Dr. Spencer's research interests include novel pharmacotherapies in children, adolescents, and adults with ADHD, tic disorders or Tourette syndrome, as well as family genetic studies of subtypes of ADHD and associated comorbidity. He has served as principal investigator and co-investigator for more than 25 studies devoted to these and other related topics. In addition, he has co-authored 95 articles, and his abstracts, book chapters, and presentations in national scientific meetings collectively number more than 140.

Norman Sussman, MD, is associate professor at the University of California at Irvine. In addition, she is a psychologist in the Department of Pediatrics at the Child Development Center and a visiting lecturer for the School of Social Sciences, also at the University of California, Irvine. Dr. Wigal's research interests include the study and treatment of ADHD, particularly in children. She has authored and co-authored numerous articles and book chapters concerning the stimulant medications for the treatment of ADHD and the link between the dopamine D4 receptor gene and ADHD, among others. Dr. Wigal is the recipient of the New Investigator Award for the New Clinical Drug Evaluation Unit, as well as the K-23 Mentored Patient-Oriented Research Career Award from the National Institute of Mental Health, which provides 5 years of funding.

Jeffrey H. Newcorn, MD, is associate professor of psychiatry and pediatrics at Mount Sinai School of Medicine, and director of the Division of Child and Adolescent Psychiatry at Mount Sinai Hospital, both in New York City. Dr. Newcorn's clinical and research interests include the study and treatment of children and adolescents with ADHD, the relationship between central serotonergic function and aggression in prepubertal boys, and the impairment and deportment responses to varying methylphenidate doses in children with ADHD.

MentalFitness. 2002;1(1):82a-l

ABSTRACT

A growing number of antidepressants are on the market. In addition, the majority of the prescriptions for antidepressants in the United States are currently written by primary care practitioners and by a growing number of obstetrician-gynecologists. Many physicians rely on data from clinical trials when making treatment choices for their patients. However, much of the information about drug efficacy can be confusing and, at times, misleading. Clinicians must learn to critically analyze the information presented in medical literature.

The latest research suggests that, contrary to conventional wisdom, there may be clinically significant differences in efficacy between various groups of antidepressants, and that many of the randomized controlled trials (RCTs) used to assess them may simply not have the statistical power to detect these differences. Since there are limitations to what conclusions can be drawn from individual RCTs, statistical methods that combine data from several studies may be used as more sensitive tools for comparative drug analysis.

Jan A. Fawcett, MD

Introduction

Randomized controlled trials (RCTs) are widely used to obtain clinical information about the effectiveness of various medications and to compare drugs with each other. Although the RCT is one of the best methods currently available, there are limitations to what conclusions can be drawn from these studies. It is important for clinicians to know and understand what these limitations are, in order to evaluate and interpret the findings of clinical studies that are presented to us in the literature.

Nine new antidepressants have been approved by the Food and Drug Administration (FDA) since 1987, and many more are being developed. Despite the growing number of antidepressants available on the market, no single agent or group of agents has been determined to be of superior efficacy compared to others. It is generally thought that all FDA-approved antidepressants are more or less equally effective. However, some evidence has suggested that clinically significant differences in efficacy do exist between various groups of antidepressants, and that many of the RCTs may not have the statistical power to detect such differences.

A recent article by Michael E. Thase, MD, in *Psychopharmacology Bulletin*[1] focused on the difficulties and limitations associated with drawing conclusions about differential efficacy of antidepressants from RCTs. It also reviewed various analytical methods used to compile data from multiple clinical studies, and compared the outcomes, advantages, and limitations of these methods. Thase also used an example of his own published pooled analysis to illustrate the potential benefits, but also the criticisms that can be leveled at the conclusions of such an analysis. In this monograph the major points raised by Thase and their relevance to the primary care physicians are reviewed.

Characteristics and Limitations of RCTs of Antidepressants

Most clinical trials of antidepressants are designed as parallel group, double-blind studies. Earlier trials typically compare the study drug to a placebo to demonstrate that its effect is superior to that of placebo. Later in the process, studies that compare the novel treatment to an established treatment are also conducted. Including standard drugs as a comparators provides information about the new drug's relative efficacy. The majority of RCTs of new antidepressants are sponsored by the manufacturer of the drug and their main objective is to rapidly obtain FDA approval for the medication. While it is necessary to provide evidence of the study drug's superiority to placebo (FDA approval requires proof of efficacy from at least two placebo-controlled trials), it is neither necessary nor advantageous for the manufacturer of the new drug to conduct definitive comparative studies with established antidepressants before the new drug passes regulatory review. This is why only limited data on comparative efficacy are available for each drug at the time of FDA approval.

A review of published RCTs has shown that all presently available antidepressants have a fairly small pharmacologic effect—only about a third of the therapeutic benefit of these drugs can be attributed to active pharmacologic effect. It might surprise many that the pharmacologic activity of antidepressants is so small, but it shows how important the placebo effect is in some of the improvement we see in patients. We know that placebo effects are common in all areas of medicine, and more recently there have been studies suggesting that there is a biological or biochemical basis for this response. In my opinion, placebo effect should not be looked down on in any way, because it is a real response, and the best physicians are those who are able to enhance the placebo response.

The ability of a study to show statistical difference depends on several factors, including effect size and sample size. When two treatments are compared, the effect size can be seen as a mean difference or standard deviation of the differences in effectiveness. Statistical power indicates the likelihood of observing a predicted effect size under predetermined conditions. Relatively small effect sizes are common for RCTs of antidepressants. It is very difficult to detect modest differences between two treatments when a small number of subjects are included in a study. Large numbers of observations are needed before one can be confident that a real difference exists. For example, if two treatments are compared, a better antidepressant may only increase the response rates by as little as 5% to 10%. An RCT would need to include at least 300 patients per arm to have enough statistical power to reliably detect such a small difference between two effective treatments. Since many antidepressant studies include fewer study subjects, they do not have the power to uncover these differences. Many of the RCTs of antidepressants are also unable to find significant differences between drug and placebo. The average drug/placebo difference in antidepressant studies is small and trials that enroll insufficient numbers of patients do not have enough statistical power to detect these modest differences.

A number of other factors can bias the outcomes

Jan A. Fawcett, MD

of RCTs. Patient-selection criteria can greatly influence the outcome of the study. For example, the proportion of treatment-resistant patients in the study group, severity of the condition, age and gender of the patients, and other characteristics may actually change the outcome of some of these studies. It is also important to question whether a study was implemented fairly—whether equally effective doses of drugs were used in different study groups.

Flaws in the methodology of analysis can also significantly impact study outcomes. For example, it is important to consider how to handle the data from patients who drop out of studies prematurely. An analysis that only includes completers as opposed to all of those who originally went into treatment can overestimate benefits, because dropouts usually have poorer outcomes than completers. On the other hand, intent-to-treat analysis, which counts the outcomes of all patients that begin treatment, is somewhat negatively biased, because the general tendency of many depressed patients to get better over time is not taken into consideration.

Quantitative Meta-Analysis Versus Pooled Analysis

Statistical methods that combine and evaluate data from several studies can be used to compare different treatment strategies. Quantitative meta-analysis and pooled analysis are two such techniques, each with their own benefits and limitations. Quantitative meta-analysis works well when a large number of relevant studies are available, because each study is considered as one unit of observation. This method allows trends in smaller studies to be assessed and various factors that can possibly influence effect size can be examined. However, it has relatively small statistical power if not many studies are available for comparison.

Pooled analysis, which combines raw original data from a set of relevant studies, considers each patient's outcome as one unit of observation, which gives this analysis a great deal more statistical power. This is an important advantage, because it allows the significance of various interactions (ie, between treatments and time or patient characteristics) to be evaluated, interactions that RCTs or other meta-analytic techniques may not have the power to detect. Results of a pooled analysis may be skewed by the larger trials included in the study, but the validity of the findings can be validated by repeating the calculations after removing each study.

Several factors such as publication bias and study selection criteria can significantly affect the outcome of these analytical methods. If only published RCTs with positive outcomes are included in an analysis, and unpublished or "failed" studies with negative results are not available or are omitted, then the conclusions that are drawn from these analyses can be misleading. The studies for comparative analysis must be pooled based on similar characteristics and must be representative. Results of these studies are largely dependent on the quality of individual studies, and are sensitive to exclusion and inclusion of individual studies.

Conclusion

The results of clinical trials can significantly affect the day-to-day practice of medical professionals. Therefore, all clinicians must learn to view the findings of clinical studies, as well as information presented in medical literature in general, with a critical eye. Although the RCT is one of the best methods currently available for evaluating the effectiveness of medical interventions, there are a number of limitations related to its methodology and characteristics. Many RCTs of antidepressants lack the statistical power to uncover small differences in efficacy that may exist between the older and newer classes of antidepressants. Some of the factors that contribute to the problem of underpowered studies with antidepressants are small effect sizes that are characteristic for antidepressants in placebo-controlled trials, and the use of insufficient sample sizes. At least 300 patients need to be enrolled in these RCTs in order to reliably detect small differences in efficacy between antidepressants. In addition, since the manufacturers of the new drug are not required to provide definitive evidence of efficacy in relation to other antidepressants, not many of these studies are conducted before the new agent obtains FDA approval.

Methods that combine data from a number of related studies can be used to overcome some of the limitations of individual trials. If clinically meaningful differences do exist between various antidepressants, it is important to show them. Quantitative meta-analyses and pooled analyses may help to uncover the differences in efficacy of various agents that individual RCTs may not have the power to show. We hope that in the future better antidepressants are developed that have more therapeutic effect than the agents we currently have available, as well as better methods to evaluate and compare the different treatment options. MF

Norman Sussman, MD

Introduction

Why is it important to understand the problems of comparative drug analysis and what is the relevance of the issues raised in Michael Thase's article to primary care physicians? Most of the prescriptions for antidepressants in the United States are currently written by primary care physicians, who are also the focus of heavy promotional activity by the pharmaceutical industry through which messages are communicated about both effectiveness and tolerability of these agents. Therefore, given the role that this group of doctors plays in the system of delivering care for people with depression and other conditions treated with antidepressants, primary care physicians need to know whether the information they are being given is accurate or whether it has been manipulated.

Interpreting the Results of RCTs

From any standard textbook we can typically read that RCTs have shown that all antidepressants are roughly equivalent in their effectiveness. At first this seems like a very simple statement, but we have to take a closer look at how the studies of antidepressants are designed and conducted, and what they can tell us. We have to view with some skepticism how well these studies correlate with what we actually see in clinical practice.

Randomized placebo-controlled trials usually have so many exclusion and inclusion criteria that the people who are included in them often bear very little resemblance to the types of people we treat in our psychiatric practices, and certainly to the types of patients primary care physicians encounter. For example, primary care physicians often see people who have comorbid medical disorders and may have comorbid psychiatric disorders, as well as patients who are taking multiple medications. Any of these characteristics would exclude a person from participating in a randomized placebo-controlled trial. How can we then apply the findings of these trials to our patients?

Another part of this problem is that it is very expensive and difficult to conduct these clinical trials and before starting a study, companies generally have their statisticians figure out how many people they need to enroll in order to come up with statistically meaningful findings. They usually do not go beyond what they think is appropriate, because it costs thousands of dollars per each evaluable patient. The goal of most studies is to demonstrate equivalence of the new agent to prevailing treatments, to standard treatments, and these studies are usually not set up or powered to demonstrate small differences between drugs. It is remarkable if these differences appear, but generally that is not the case.

It is also important to remember that one study does not conclusively prove anything. A single clinical trial being presented in either an academic or a commercial setting can be very misleading if it suggests that either a difference exists or does not exist between two different treatment options. There are many studies that do not get published at all for a variety of reasons, the most common being that they are considered "failed studies." For example, during the course of the study it may be found that placebo response rate is very high, or some other factor becomes apparent that the sponsor of the study does not want to have published, because it does not reflect well on their product. There is no obligation to publish a study just because it has been completed, and if only studies with positive results are made public we are not getting the complete picture.

Results Across Studies

Given the limitations of individual trials, the idea of using methods that pool information from several studies is widely accepted in other areas of medicine outside of psychiatry. For example, such analyses have been done with cholesterol-lowering drugs statins, with angiotensin-converting enzyme (ACE) inhibitors, and other drugs. It is a highly respected way of looking for differences in situations where large numbers of patients need to be treated in order for small differences to emerge.

A pooled analysis by Thase and colleagues[2] has raised this issue in psychiatry. In that analysis, the treatment outcomes of over 2,000 patients were pooled. Patients treated with venlafaxine, a dual-acting antidepressant affecting both norepinephrine and serotonin systems, were compared with patients treated with selective serotonin reuptake inhibitors (SSRIs), mainly fluoxetine. The results of this analysis showed that there was about a 10% difference between the two treatment groups in terms of patients achieving remission (defined as a total score of 7 or less on the Hamilton Rating Scale for Depression). About 45% of the patients treated with venlafaxine achieved remission, while about 35% of those treated with SSRIs had the same outcome. This is a meaningful difference. In a primary care setting, for example, a 10% difference between two ACE inhibitors, or cholesterol-lowering drugs, or even antibiotics, in terms of people completely recovering as opposed to simply showing some

improvement, would certainly be considered significant. If such differences exist between antidepressants, it is obviously important to find them.

This pooled analysis and other similar studies have stirred up quite a bit of controversy and interest in whether or not the way we look at measuring drug activity and drug effectiveness is misleading or whether it is informative. The companies that market SSRIs have been picking apart the validity of doing such an analysis.

One important aspect that makes the pooled analysis by Thase and colleagues different from many other pooled analyses is that it included all studies, published and unpublished, that were available at the time comparing venlafaxine to SSRIs, without selectively choosing based on the findings and making the result look better than it is. There are other pooled analyses that have also suggested high remission rates, but more critical review would uncover several shortcomings. For example, studies of different design and methodology have been pooled. It makes a big difference whether a study is open-label or double-blind and placebo-controlled. Open-label studies are likely to produce higher remission or response rates and cannot be combined with double-blind studies in comparative drug analysis.

Duration of treatment is also significant, because treatment outcomes of patients who have been in therapy for 6 months cannot be compared with those of patients who have been treated for only 8 weeks, as they are two completely different populations. Clearly, over time more people are going to achieve remission than with treatment of shorter duration.

Another question that has been raised about doing a pooled analysis is how do we know that in the various studies comparably effective or equivalent doses of different drugs were used? For example, nobody knows whether 150 mg of venlafaxine is equal to 20, 30, or 40 mg of paroxetine or fluoxetine. Most of us consider 20–40 mg of an SSRI to be a very typical dose and in primary care these are the doses that are used. Similarly, 150 mg of venlafaxine is considered to be a reasonably effective dose.

Some aspects of the pooled analysis done by Thase's group could also be criticized. For example, if a large number of patients included in the studies pooled in Thase's analysis had previously failed treatment with other SSRIs, the greater efficacy of venlafaxine that was shown could be accounted for by the SSRI-resistant patients. Also, since majority of the patients in the SSRI group were treated with fluoxetine, it cannot be conclusively stated that venlafaxine is more efficacious than all SSRIs.

However, one of the benefits of doing a large pooled analysis is that it controls for many minor variations and fluctuations. This is why it is important to stress that although pooled analysis has its limitations, it is a legitimate technique and, when done properly, can show meaningful differences.

Conclusion

Analytical methods that combine results from many studies are good tools for comparing different treatment strategies. They have more statistical power to detect small but important differences that individual RCTs may not be able to find. We also have to better educate and inform clinicians about how clinical trials are conducted and what their limitations are, because often information is presented to us that, at first glance, may look very impressive, and we have to learn to critically evaluate these findings. It is my only hope that clinicians care enough to want to understand these issues and become informed consumers of RCT findings. MF

References

1. Thase ME. Comparing the methods used to compare antidepressants. *Psychopharm Bull*. 2002;36[Suppl 1]:4-17.
2. Thase ME, Entsuah AR, Rudolph RL. Remission rates during treatment with venlafaxine or selective serotonin reuptake inhibitors. *Br J Psychiatry*. 2001;178:234-241.

Jeffrey H. Newcorn, MD

Introduction

Randomized controlled trials (RCTs) are widely used to obtain clinical information about the effectiveness of various medications and to compare drugs with each other. Although the RCT is one of the best methods currently available, there are limitations to what conclusions can be drawn from these studies. It is important for clinicians to know and understand what these limitations are, in order to evaluate and interpret the findings of clinical studies that are presented to us in the literature.

Nine new antidepressants have been approved by the Food and Drug Administration (FDA) since 1987, and many more are being developed. Despite the growing number of antidepressants available on the market, no single agent or group of agents has been determined to be of superior efficacy compared to others. It is generally thought that all FDA-approved antidepressants are more or less equally effective. However, some evidence has suggested that clinically significant differences in efficacy do exist between various groups of antidepressants, and that many of the RCTs may not have the statistical power to detect such differences.

A recent article by Michael E. Thase, MD, in *Psychopharmacology Bulletin*[1] focused on the difficulties and limitations associated with drawing conclusions about differential efficacy of antidepressants from RCTs. It also reviewed various analytical methods used to compile data from multiple clinical studies, and compared the outcomes, advantages, and limitations of these methods. Thase also used an example of his own published pooled analysis to illustrate the potential benefits, but also the criticisms that can be leveled at the conclusions of such an analysis. In this monograph the major points raised by Thase and their relevance to the primary care physicians are reviewed.

Characteristics and Limitations of RCTs of Antidepressants

Most clinical trials of antidepressants are designed as parallel group, double-blind studies. Earlier trials typically compare the study drug to a placebo to demonstrate that its effect is superior to that of placebo. Later in the process, studies that compare the novel treatment to an established treatment are also conducted. Including standard drugs as a comparators provides information about the new drug's relative efficacy. The majority of RCTs of new antidepressants are sponsored by the manufacturer of the drug and their main objective is to rapidly obtain FDA approval for the medication. While it is necessary to provide evidence of the study drug's superiority to placebo (FDA approval requires proof of efficacy from at least two placebo-controlled trials), it is neither necessary nor advantageous for the manufacturer of the new drug to conduct definitive comparative studies with established antidepressants before the new drug passes regulatory review. This is why only limited data on comparative efficacy are available for each drug at the time of FDA approval.

A review of published RCTs has shown that all presently available antidepressants have a fairly small pharmacologic effect—only about a third of the therapeutic benefit of these drugs can be attributed to active pharmacologic effect. It might surprise many that the pharmacologic activity of antidepressants is so small, but it shows how important the placebo effect is in some of the improvement we see in patients. We know that placebo effects are common in all areas of medicine, and more recently there have been studies suggesting that there is a biological or biochemical basis for this response. In my opinion, placebo effect should not be looked down on in any way, because it is a real response, and the best physicians are those who are able to enhance the placebo response.

The ability of a study to show statistical difference depends on several factors, including effect size and sample size. When two treatments are compared, the effect size can be seen as a mean difference or standard deviation of the differences in effectiveness. Statistical power indicates the likelihood of observing a predicted effect size under predetermined conditions. Relatively small effect sizes are common for RCTs of antidepressants. It is very difficult to detect modest differences between two treatments when a small number of subjects are included in a study. Large numbers of observations are needed before one can be confident that a real difference exists. For example, if two treatments are compared, a better antidepressant may only increase the response rates by as little as 5% to 10%. An RCT would need to include at least 300 patients per arm to have enough statistical power to reliably detect such a small difference between two effective treatments. Since many antidepressant studies include fewer study subjects, they do not have the power to uncover these differences. Many of the RCTs of antidepressants are also unable to find significant differences between drug and placebo. The average drug/placebo difference in antidepressant studies is small and trials that enroll insufficient numbers of patients do not have enough statistical power to detect these modest differences.

A number of other factors can bias the outcomes

Jeffrey H. Newcorn, MD

of RCTs. Patient-selection criteria can greatly influence the outcome of the study. For example, the proportion of treatment-resistant patients in the study group, severity of the condition, age and gender of the patients, and other characteristics may actually change the outcome of some of these studies. It is also important to question whether a study was implemented fairly—whether equally effective doses of drugs were used in different study groups.

Flaws in the methodology of analysis can also significantly impact study outcomes. For example, it is important to consider how to handle the data from patients who drop out of studies prematurely. An analysis that only includes completers as opposed to all of those who originally went into treatment can overestimate benefits, because dropouts usually have poorer outcomes than completers. On the other hand, intent-to-treat analysis, which counts the outcomes of all patients that begin treatment, is somewhat negatively biased, because the general tendency of many depressed patients to get better over time is not taken into consideration.

Quantitative Meta-Analysis Versus Pooled Analysis

Statistical methods that combine and evaluate data from several studies can be used to compare different treatment strategies. Quantitative meta-analysis and pooled analysis are two such techniques, each with their own benefits and limitations. Quantitative meta-analysis works well when a large number of relevant studies are available, because each study is considered as one unit of observation. This method allows trends in smaller studies to be assessed and various factors that can possibly influence effect size can be examined. However, it has relatively small statistical power if not many studies are available for comparison.

Pooled analysis, which combines raw original data from a set of relevant studies, considers each patient's outcome as one unit of observation, which gives this analysis a great deal more statistical power. This is an important advantage, because it allows the significance of various interactions (ie, between treatments and time or patient characteristics) to be evaluated, interactions that RCTs or other meta-analytic techniques may not have the power to detect. Results of a pooled analysis may be skewed by the larger trials included in the study, but the validity of the findings can be validated by repeating the calculations after removing each study.

Several factors such as publication bias and study selection criteria can significantly affect the outcome of these analytical methods. If only published RCTs with positive outcomes are included in an analysis, and unpublished or "failed" studies with negative results are not available or are omitted, then the conclusions that are drawn from these analyses can be misleading. The studies for comparative analysis must be pooled based on similar characteristics and must be representative. Results of these studies are largely dependent on the quality of individual studies, and are sensitive to exclusion and inclusion of individual studies.

Conclusion

The results of clinical trials can significantly affect the day-to-day practice of medical professionals. Therefore, all clinicians must learn to view the findings of clinical studies, as well as information presented in medical literature in general, with a critical eye. Although the RCT is one of the best methods currently available for evaluating the effectiveness of medical interventions, there are a number of limitations related to its methodology and characteristics. Many RCTs of antidepressants lack the statistical power to uncover small differences in efficacy that may exist between the older and newer classes of antidepressants. Some of the factors that contribute to the problem of underpowered studies with antidepressants are small effect sizes that are characteristic for antidepressants in placebo-controlled trials, and the use of insufficient sample sizes. At least 300 patients need to be enrolled in these RCTs in order to reliably detect small differences in efficacy between antidepressants. In addition, since the manufacturers of the new drug are not required to provide definitive evidence of efficacy in relation to other antidepressants, not many of these studies are conducted before the new agent obtains FDA approval.

Methods that combine data from a number of related studies can be used to overcome some of the limitations of individual trials. If clinically meaningful differences do exist between various antidepressants, it is important to show them. Quantitative meta-analyses and pooled analyses may help to uncover the differences in efficacy of various agents that individual RCTs may not have the power to show. We hope that in the future better antidepressants are developed that have more therapeutic effect than the agents we currently have available, as well as better methods to evaluate and compare the different treatment options. MF

Lecture Slide Series

The Evolution of Antidepressants

1.MF_TM1

1950s: Tricyclic antidepressants

1960s: Monoamine oxidase inhibitors

1970s: Heterocyclic antidepressants

1980s: Selective serotonin reuptake inhibitors

1990s: Receptor modulators, selective norepinephrine reuptake inhibitors, dual reuptake inhibitors

Source: Kaplan HI, Sadock BJ. *Pocket Handbook of Psychiatric Drug Treatment.* 2nd ed. Baltimore, MD: Williams & Wilkins; 1996.
Fawcett JA, Sussman N. *MentalFitness*. Vol 1, No 1. 2002.

Critical Elements of RCTs

2.MF_TM1

- Prospective observation of outcomes
- Reliable and valid assessment of outcomes
- Treatment administered according to standardized protocol
- Randomization to study treatments to help ensure comparable study groups
- Hypotheses and principal statistical tests determined in advance
- Outcomes assessed without knowledge of treatment assignment (double-blind) to minimize effects of expectations

RCTs=randomized controlled trials.
Source: Thase ME. *Psychopharm Bull*. 2002;36(Suppl 1):4-17.
Fawcett JA, Sussman N. *MentalFitness*. Vol 1, No 1. 2002.

RCTs of Antidepressants

3.MF_TM1

- RCTs of antidepressants typically use parallel group, double-blind design
- Studies can last from 4 to 12 weeks
- Early clinical trials typically compare the study drug to placebo (2-arm studies)
- Later clinical trials compare the novel treatment to placebo and an established treatment (3-arm studies)
- After FDA approval, studies comparing the novel antidepressant to an established comparator (without a placebo group) are also conducted

RCTs=randomized controlled trials; FDA=Food and Drug Administration.
Fawcett JA, Sussman N. *MentalFitness*. Vol 1, No 1. 2002.

RCTs of Antidepressants (cont.)

4.MF_TM1

- Pharmacologic effect of antidepressants is small (about 1/3 of therapeutic benefit)
- Modest effect sizes are common for RCTs of antidepressants
- RCTs that include less than 300 study subjects per arm do not have the statistical power to detect small differences between active treatments

RCTs=randomized controlled trials.
Fawcett JA, Sussman N. *MentalFitness*. Vol 1, No 1. 2002.

Open-Label Versus Double-Blind, Placebo-Controlled Studies

5.MF_TM1

- Open-label studies
  - Likely to produce higher response or remission rates, because there is an expectation or anticipation of treatment
  - May overrate efficacy of particular agents
  - Provide the first level of groundwork to determine if there is evidence of significant clinical efficacy for a particular agent
- Double-blind, placebo-controlled studies
  - Should follow if open-label studies provide positive results
  - Are necessary for comparative drug analysis
  - Different forms of un-blinding may also lead to overestimation of drug efficacy

Fawcett JA, Sussman N. *MentalFitness*. Vol 1, No 1. 2002.

Duration of Treatment or Follow-Up in Evaluating Treatment Outcomes

6.MF_TM1

- Studies of antidepressant efficacy in which the duration of treatment or follow-up vary significantly cannot be directly compared
- Longer treatment or follow-up will inevitably result in better rates of remission
- Proportion of patients who are incompletely remitted after a few weeks or months of therapy may become fully remitted after longer period of therapy

Fawcett JA, Sussman N. *MentalFitness*. Vol 1, No 1. 2002.

Lecture Slide Series

The Role of Premature Discontinuation of Treatment in RCT Analysis

7.MF_TM1

- 20% to 40% of subjects enrolled in clinical studies discontinue treatment prematurely
- Analysis of study completers can overestimate treatment outcomes (dropouts generally have poorer outcomes than completers)
- Intent-to-treat analysis counts the outcomes of all patients who begin treatment, but can be somewhat negatively biased (the general tendency of patients to get better over time is not taken into consideration)

RCT=randomized controlled trial..
Fawcett JA, Sussman N. *MentalFitness*. Vol 1, No 1. 2002.

Placebo Response in Antidepressant Studies

8.MF_TM1

- Placebo effects in depression studies vary greatly depending on the severity of illness (from close to 0% in studies of severe depression to over 50% in studies of mild depression)
- The most efficient way to evaluate antidepressant efficacy is within a group of patients with moderate to moderate-severe depression
- The propensity towards high placebo rates and issues in blinding of treatment conditions in depression studies has led to much criticism and skepticism about antidepressant trials

Sources: Thase ME. *J Clin Psychiatry*. 1999;60:23-31; Quitkin FM. *Am J Psychiatry*. 1999; 156:829-836.
Fawcett JA, Sussman N. *MentalFitness*. Vol 1, No 1. 2002.

Characteristics and Advantages of Pooled Analysis

9.MF_TM1

- Combines raw original data from each study
- Unit of observation (ie, “N”) is each individual subject from the data set
- Large N group increases statistical power when small number of studies are available
- Examines correlates of response
- Allows the analysis of data from specific subgroups

Sources: Thase ME. *J Clin Psychiatry*. 1999;60:23-31; Thase ME, Entsuah AR, Rudolph RL. *Br J Psychiatry*. 2001;178:234-241; Moncrieff J, Wessely S, Hardy R. *Br J Psychiatry*. 1998;172:227-231; Gilbody SM, Song F, Eastwood AJ, Sutton A. *Acta Psychiatr Scand*. 2000;102:241-249.
Fawcett JA, Sussman N. *MentalFitness*. Vol 1, No 1. 2002.

Characteristics and Advantages of Quantitative Meta-Analysis

10.MF_TM1

- Useful when large number of studies are available
- Unit of observation (ie, “N”) is each study
- Increases statistical power of drug efficacy
- Allows trends in small studies to be reliably assessed
- Able to examine factors that may influence of moderate effect sizes

Sources: Thase ME. *J Clin Psychiatry*. 1999;60:23-31; Thase ME, Entsuah AR, Rudolph RL. *Br J Psychiatry*. 2001;178:234-241; Moncrieff J, Wessely S, Hardy R. *Br J Psychiatry*. 1998;172:227-231; Gilbody SM, Song F, Eastwood AJ, Sutton A. *Acta Psychiatr Scand*. 2000;102:241-249.
Fawcett JA, Sussman N. *MentalFitness*. Vol 1, No 1. 2002.

Disadvantages of Pooled Analysis and Quantitative Meta-Analysis

11.MF_TM1

Disadvantages of both methods:
- Conclusions can be misleading because of publication bias
- Limited by bias of selection of non-representative studies
- Results may be skewed by the quality of individual studies
- Sensitive to inclusion and exclusion of individual studies

Disadvantages of pooled analysis:
- Contingent on the reliability of dependent measures
- Results may be skewed by the larger studies in the analysis

Disadvantage of meta-analysis:
- Small N group=small power compared to pooled analysis

Sources: Thase ME. *J Clin Psychiatry*. 1999;60:23-31; Thase ME, Entsuah AR, Rudolph RL. *Br J Psychiatry*. 2001;178:234-241; Moncrieff J, Wessely S, Hardy R. *Br J Psychiatry*. 1998;172:227-231; Gilbody SM, Song F, Eastwood AJ, Sutton A. *Acta Psychiatr Scand*. 2000;102:241-249.
Fawcett JA, Sussman N. *MentalFitness*. Vol 1, No 1. 2002.

Conclusions: RCTs of Antidepressants

12.MF_TM1

- Most RCTs of antidepressants do not have the statistical power to detect small differences in efficacy between active treatments
- The literature on comparative drug efficacy should be analyzed critically, keeping in mind differences and limitations of study design and methods
- Statistical methods that combine data from a related set of studies (ie, quantitative meta-analysis and pooled analysis) can be used as more sensitive tools for comparative drug analysis

RCTs=randomized controlled trials.
Fawcett JA, Sussman N. *MentalFitness*. Vol 1, No 1. 2002.

Lecture Slide Series

The Role of Premature Discontinuation of Treatment in RCT Analysis

7.MF_TM1

- 20% to 40% of subjects enrolled in clinical studies discontinue treatment prematurely
- Analysis of study completers can overestimate treatment outcomes (dropouts generally have poorer outcomes than completers)
- Intent-to-treat analysis counts the outcomes of all patients who begin treatment, but can be somewhat negatively biased (the general tendency of patients to get better over time is not taken into consideration)

RCT=randomized controlled trial..
Fawcett JA, Sussman N. *MentalFitness*. Vol 1, No 1. 2002.

Placebo Response in Antidepressant Studies

8.MF_TM1

- Placebo effects in depression studies vary greatly depending on the severity of illness (from close to 0% in studies of severe depression to over 50% in studies of mild depression)
- The most efficient way to evaluate antidepressant efficacy is within a group of patients with moderate to moderate-severe depression
- The propensity towards high placebo rates and issues in blinding of treatment conditions in depression studies has led to much criticism and skepticism about antidepressant trials

Sources: Thase ME. *J Clin Psychiatry*. 1999;60:23-31; Quitkin FM. *Am J Psychiatry*. 1999; 156:829-836.
Fawcett JA, Sussman N. *MentalFitness*. Vol 1, No 1. 2002.

Characteristics and Advantages of Pooled Analysis

9.MF_TM1

- Combines raw original data from each study
- Unit of observation (ie, “N”) is each individual subject from the data set
- Large N group increases statistical power when small number of studies are available
- Examines correlates of response
- Allows the analysis of data from specific subgroups

Sources: Thase ME. *J Clin Psychiatry*. 1999;60:23-31; Thase ME, Entsuah AR, Rudolph RL. *Br J Psychiatry*. 2001;178:234-241; Moncrieff J, Wessely S, Hardy R. *Br J Psychiatry*. 1998;172:227-231; Gilbody SM, Song F, Eastwood AJ, Sutton A. *Acta Psychiatr Scand*. 2000;102:241-249.
Fawcett JA, Sussman N. *MentalFitness*. Vol 1, No 1. 2002.

Characteristics and Advantages of Quantitative Meta-Analysis

10.MF_TM1

- Useful when large number of studies are available
- Unit of observation (ie, “N”) is each study
- Increases statistical power of drug efficacy
- Allows trends in small studies to be reliably assessed
- Able to examine factors that may influence of moderate effect sizes

Sources: Thase ME. *J Clin Psychiatry*. 1999;60:23-31; Thase ME, Entsuah AR, Rudolph RL. *Br J Psychiatry*. 2001;178:234-241; Moncrieff J, Wessely S, Hardy R. *Br J Psychiatry*. 1998;172:227-231; Gilbody SM, Song F, Eastwood AJ, Sutton A. *Acta Psychiatr Scand*. 2000;102:241-249.
Fawcett JA, Sussman N. *MentalFitness*. Vol 1, No 1. 2002.

Disadvantages of Pooled Analysis and Quantitative Meta-Analysis

11.MF_TM1

Disadvantages of both methods:
- Conclusions can be misleading because of publication bias
- Limited by bias of selection of non-representative studies
- Results may be skewed by the quality of individual studies
- Sensitive to inclusion and exclusion of individual studies

Disadvantages of pooled analysis:
- Contingent on the reliability of dependent measures
- Results may be skewed by the larger studies in the analysis

Disadvantage of meta-analysis:
- Small N group=small power compared to pooled analysis

Sources: Thase ME. *J Clin Psychiatry*. 1999;60:23-31; Thase ME, Entsuah AR, Rudolph RL. *Br J Psychiatry*. 2001;178:234-241; Moncrieff J, Wessely S, Hardy R. *Br J Psychiatry*. 1998;172:227-231; Gilbody SM, Song F, Eastwood AJ, Sutton A. *Acta Psychiatr Scand*. 2000;102:241-249.
Fawcett JA, Sussman N. *MentalFitness*. Vol 1, No 1. 2002.

Conclusions: RCTs of Antidepressants

12.MF_TM1

- Most RCTs of antidepressants do not have the statistical power to detect small differences in efficacy between active treatments
- The literature on comparative drug efficacy should be analyzed critically, keeping in mind differences and limitations of study design and methods
- Statistical methods that combine data from a related set of studies (ie, quantitative meta-analysis and pooled analysis) can be used as more sensitive tools for comparative drug analysis

RCTs=randomized controlled trials.
Fawcett JA, Sussman N. *MentalFitness*. Vol 1, No 1. 2002.