CENTER FOR DRUG EVALUATION AND RESEARCH

Guidance for Industry

The FDA published Good Guidance Practices in February 1997. This guidance was developed and issued prior to that date.

Additional copies are available from:
Office of Training and Communications
Division of Communications Management
Drug Information Branch, HFD-210
5600 Fishers Lane
Rockville, MD 20857

(Tel) 301-827-4573
(Internet) http://www.fda.gov/cder/guidance/index.htm

U.S. DEPARTMENT OF HEALTH AND HUMAN SERVICES, FOOD AND DRUG ADMINISTRATION

guidelines for the clinical evaluation of Antianxiety Drugs

U.S. DEPARTMENT OF HEALTH, EDUCATION, AND WELFARE
Public Health Service
Food and Drug Administration

HEW (FDA) 77-3043

GUIDELINES FOR THE CLINICAL EVALUATION OF ANTIANXIETY DRUGS

September 1977

Comments on the contents of this publication are invited and should be addressed to the following office:

Group Leader for Psychopharmacological Drugs
Division of Neuropharmacological Drug Products
Bureau of Drugs
Food and Drug Administration
5600 Fishers Lane
Rockville, Maryland 20857

ABSTRACT

The Food and Drug Administration, with the assistance of its scientific Advisory Committees and other outside consultants, the American Academy of Pediatrics' Committee on Drugs, and consultants to the Pharmaceutical Manufacturers' Association has developed guidelines for the clinical evaluation of new drugs. These guidelines present acceptable current approaches to the study of investigational drugs in man, and pertain to Phases I through III of the investigation. They represent generally accepted principles for arriving at valid conclusions concerning safety and effectiveness of new drugs, as well as the views of outstanding experts concerning appropriate methods of study of specific classes of drugs.

The FDA welcomes comments on the guidelines, and expects to keep them current by review and update at approximately two-year intervals.

FOREWORD

The purpose of these guidelines is to present acceptable current approaches to the study of investigational drugs in man. These guidelines contain both generalities and specifics and were developed from experience with available drugs. It is anticipated that with the passage of time these guidelines will require revision. In order to keep them current a re-review will be performed approximately every 18 to 24 months.

These guidelines are not to be interpreted as mandatory requirements by the FDA to allow continuation of clinical trials with investigational drugs or to obtain approval of a new drug for marketing. These guidelines, in part, contain recommendations for clinical studies which are recognized as desirable approaches to be used in arriving at conclusions concerning safety and effectiveness of new drugs; and in the other part they consist of the views of outstanding experts in the field as to what constitutes appropriate methods of study of specific classes of drugs. In some cases other methods may be equally applicable or newer methods may be preferable, and for certain entirely new entities it is possible that the guidelines may be only minimally applicable.

Under FDA regulations (21 CFR 10.90(b)) all clinical guidelines constitute advisory opinions on an acceptable approach to meeting regulatory requirements, and research begun in good faith under such guidelines will be accepted by the Agency for review purposes unless this guideline (or the relevant portion of it) has been formally rescinded for valid health reasons. This does not imply that results obtained in studies conducted under these guidelines will necessarily result in the approval of an application or that the studies suggested will produce the total clinical information required for approval of a particular drug.

Many of the clinical guidelines have been developed largely, or entirely, by FDA's Advisory Committees and consultants. Others were originally developed by intramural committees and consultants of FDA and of the Pharmaceutical Manufacturers Association; in these cases the guidelines were reviewed and revised, as appropriate, by FDA's Advisory Committees.

The general guidelines for the evaluation of drugs in infants and children and most of those for study of various drug classes in children were developed by the Committee on Drugs of the American Academy of Pediatrics (AAP). Some of the pediatric guidelines for specific classes were written by FDA's Advisory Committees. There was cross review and comment on the pediatric guidelines by both the Committee on Drugs of the AAP and FDA's Advisory Committees.

The Bureau of Drugs of the FDA wishes to thank the many individuals who devoted so much time and effort to the development of these guidelines.

J. Richard Crout, M.D.
Director
Bureau of Drugs

Marion J. Finkel, M.D.
Associate Director for
New Drug Evaluation
Bureau of Drugs

GUIDELINES FOR THE CLINICAL EVALUATION OF ANTIANXIETY DRUGS

"General Considerations for the Clinical Evaluation of Drugs" is an important companion piece and should be reviewed prior to reading these guidelines. It contains suggestions which are applicable to investigational drug studies for most classes of drugs and enables elimination of repetitious material in each of the specific guidelines.

1.0 INTRODUCTION

1.1 Purpose: Psychotropic drug guidelines have been developed to clarify the planning, monitoring, analysis, and evaluation of clinical studies of investigational new drugs with anticipated efficacy in the treatment of anxiety, depression and psychosis. These particular guidelines represent a reasonable set of considerations in planning and executing appropriate clinical studies at various stages of drug development and in documenting and evaluating the results of individual studies as well as groups of studies relevant to establishing safety and efficacy. Attention to these guidelines by investigators, drug company staff, and FDA hopefully will provide a common basis for planning, evaluating, and interpreting clinical studies and will facilitate the development of new and effective drugs. Time and energy invested in unnecessary or inappropriately designed and executed studies also should be reduced at all levels. The guidelines will be particularly helpful to staff (investigational, industrial, or governmental) not yet expert in the area of clinical psychopharmacology. More detailed discussions of clinical drug evaluation can be found in recent publications. [1,2,3]

The guidelines are not intended to be immutable nor are they to be used to stifle innovative approaches to clinical research. They are intended neither as minimal standards nor as unrealistic standards of excellence but are written with the hope of improving the quality and meaningfulnes of clinical studies designed to evaluate new psychotropic drugs. Although detailed guidelines are written only for antianxiety, antidepressant, and anti psychotic agents, the general issues discussed should be useful in planning studies of drug efficacy for other psychiatric indications.

1.2 Phases of Clinical Drug Evaluation: The guidelines specify the purposes of each of the phases of drug development and suggest characteristics of studies designed to provide the necessary information for proceeding to the next phase of study. The phases (I, II, III) are not entirely separable, and data collected in one phase will be relevant to decisions made in the next phase.

1.21 Phase I, human pharmacology, follows completion of appropriate pharmocological and toxicological studies in lower animals. A reasonably safe initial dose for use in man should be inferable from the data available. Although animal pharmacology may suggest that a new compound resembles existing drugs with known efficacy in specific psychiatric states, adequate animal models of human psychiatric illnesses are nonexistent, and the initiation of Phase I studies could be based on considerations other than specific pharmacological activity in animals (e.g., a possible effect on brain biochemistry). Phase I studies should provide evidence as to the safety, pharmacological effects, and dose-related side effects

in normal volunteers and/or psychiatric patients based on both single dose and time-limited multiple dose studies. Evidence as to drug absorption, distribution, excretion, and metabolism is also desirable. Obviously, if data documenting drug effects on mood or behavior in normal volunteers or on symptoms in patients can be obtained during Phase I studies, this is highly desirable.

1.22 Phase II studies are designed to provide reasonable evidence of clinical efficacy and usually proceed from carefully conducted open or single-blind studies in appropriate patient groups toward controlled studies designed to clearly establish efficacy in well-defined patient populations. Obviously, knowledge of the effective dose and side effects and some indication of therapeutic potential must be obtained from early Phase II studies before more formal controlled studies are worthwhile.

1.23 In Phase III, more extensive controlled studies are conducted to confirm and extend the findings in Phase II, aiming at more extensive evidence of efficacy under a wider range of patient groups and settings as well as more specific information about symptoms or patient types in which the drug is especially effective. Larger samples of patients are usually involved in the total work of this phase so as to obtain more information about the incidence of both common and rarer adverse effects. Controlled studies can provide the kinds of information needed to assess the causal relationship between the investigational drug used and the adverse somatic or psychiatric events (e.g., jaundice, dermatitis, and suicidal attempts). Uncontrolled studies of the drug in a wider range of clinical settings may be useful in assessing its safety, particularly the rarer adverse reactions.

1.3 General Methodology: The principles for "adequate and well-controlled clinical investigations" (Code of Federal Regulations, Title 21 Section 314.111(a)(5)(ii)) are applicable to the classes of drugs covered in these guidelines. Difficulty frequently will be encountered in establishing an unequivocal diagnosis, since psychiatry lacks clear, quantitative, universally accepted diagnostic criteria. Both anxiety and depression are subjective symptoms occurring in association with a variety of other types of psychopathology, and both symptoms often occur in the same patient in varying proportions. This makes it impossible to define clear, universally applicable diagnostic criteria for use in each of the three classes of drugs for which guidelines have been prepared. Thus, it is even more necessary that individual study protocols describe explicitly their operational criteria for selection (and exclusion) of patients and that each patient be characterized as to the presence and intensity of various aspects of psychopathology commonly observed in anxious, depressed or psychotic patients. Otherwise, the results of several studies cannot be compared or pooled.

Although existing assessment measures and study designs are often capable of detecting clear evidence of clinical drug efficacy, the state-of-the-art in clinical psychopharmacology still requires the development of new and better techniques. The development of new, potentially useful approaches to drug evaluation are to be encouraged by the investigators, the industry, and the FDA. Evidence of clinical efficacy derived from such new approaches will be given sufficient weight to encourage further improvements in the tools of clinical psychopharmacology. The absence of useful laboratory measures of anxiety or depression is regrettable. The absence of reliable methods for measuring blood levels of almost all existing psychoactive drugs is equally regrettable, since it suggests that absorption, distribution, metabolic, and excretion studies of future investigational agents may be very difficult to conduct. Limited studies using radioactively labelled compounds, of course, may be possible.

1.4 Monitoring: The task of monitoring the progress of an investigational new drug through the three phases rests primarily on the professional staff assigned by the drug company to the particular drug. In carrying out this role, close contact between industry monitors and clinical investigators is vitally needed as are well-planned, realistic study designs. The Food and Drug Administration also is responsible for

reviewing clinical research plans. Ideally, there should be continuing contact between the industry monitor and the FDA staff to help insure that studies will provide acceptable evidence for the decision to advance the drug to the next phase of evaluation or to final marketing. Using these guidelines, requirements can be formulated and made available to the clinical investigator which may serve to assist the industry monitor in motivating the investigator to adhere carefully to protocol requirements and data-collection procedures. On the other hand, some leeway must be provided clinical investigators to let them follow up unanticipated leads growing out of adequately designed formal studies. Only good communication between investigator, company monitor, and the FDA will enable the monitoring of an investigational drug through its several phases to be sensible, reasonably efficient, and productive of the data necessary to competently evaluate both the drug's efficacy and its safety. There must be sufficient allowance for flexible, exploratory studies to insure that a compound which might be effective in condition A is not restrictively consigned to evaluation only in condition B where it is ineffective.

1.5 Ethics: The possibility of teratogenicity frequently makes it necessary to avoid involving female patients of childbearing potential in studies of investigational drugs. Women with childbearing potential should be included only when results of required reproduction and teratologic studies in animals have been reviewed by the pharmacologist and are considered to be satisfactory. However, in some institutionalized severely ill female patients, the clinical setting provides sufficient insurance against the risk of pregnancy to make their inclusion in studies justifiable.

Investigational drugs should be administered to children and minor adolescents only with parental consent. Such studies should not be initiated until safety and efficacy studies in adults are well advanced. Studies in children of drugs likely to be used in the treatment of the hyperkinetic syndrome, or adolescent forms of adult illnesses should not be unduly delayed since the drug, once marketed, may well be used in such patients without the physician having the benefit of knowledge obtained from careful clinical studies. Patients with a history of drug dependency usually should be excluded. Since the elderly and patients with cardiovascular, hepatic, renal, or other organic diseases would also be potential recipients of these drugs, if marketed, it is desirable that they be cautiously included in Phase III studies unless contraindicated by the nature of the drug. The use of placebo in clinical trials of psychotropic drugs poses a special problem. In conditions such as depression and anxiety states, the superiority of standard existing drugs over placebo is of sufficiently modest extent to make the administration of placebo to some patients in a study entirely justifiable, particularly if there are explicit provisions for removing from the study patients whose clinical condition worsens or fails to improve in a reasonable period of time.

1.6 Data Collection and Recording: Definitive, accurate, and appropriate documentation of clinical trials is an absolute necessity if valid conclusions regarding safety and efficacy are to be made by investigators, the industry, and the FDA. A perfectly designed and executed clinical trial without adequate documentation is wasted effort. The details for adequate documentation need not be specifically defined in this brief guideline. In general, it is necessary to document (1) the samples studied and the nature of the population or populations to which the results of clinical trials may be generalized, (2) the procedures followed during the trial, (3) the "case history" of each patient, (4) the criterion measures (for psychopathology, laboratory and physical examinations, and side effects), and (5) other variables which may have an influence on the results of the trial. To the extent that previously standardized and validated forms and measures are available and appropriate to assess the populations and drugs being evaluated, these are preferred since results are then more easily interpreted. The emphasis or extent of documentation for measures of safety versus efficacy or degree of specificity regarding efficacy will obviously and appropriately vary from Phase I to Phase II to Phase III. These will be in keeping with the objectives of the phases and the hypotheses of the individual studies.

1.7 Data Presentation: Once the individual appropriate items of information have been accurately collected during the clinical trial (documentation), it then becomes necessary to organize and present the data in a form which allows results to be viewed and conclusions to be reached. For a single clinical trial or study, this is done by: (1) preparing an individual case record for each patient, (2) summarizing data according to the treatment groups (or other meaningful groupings) in tabular and graphic form (essentially showing frequencies), and (3) performing appropriate statistical inference tests which indicate whether observed results are (un) likely to have occurred by chance. The ways in which the data may be presented cannot be detailed here but should be appropriate to the measures employed and the design of the trial. The use of well-validated and documented statistical procedures enhances the interpretation of results, but novel (well-documented) approaches which enhance the understanding of the results of a trial are encouraged. It should be posible in a well-documented and well-presented study to trace an individual patient's raw data through to its contribution in arriving at a probability statement.

1.8 Interpretation: While the previous section dealt with data presentation from single clinical trials, it is usual that evidence for safety and efficacy is acquired over the course of many studies carried out over considerable periods of time and at different geographical locations. Interpretation refers to the process by which the evidence from these various studies is considered together in reaching a conclusion. This may be done by comparing the results of one study to another or, where appropriate, by actually combining or pooling the data from several studies into a larger analysis. Where efforts of this type are undertaken the reasons for this and the procedures employed should be clearly given. The "picture" resulting from pooling of data should be clearer or more comprehensive than the results from the component individual studies.

EARLY CLINICAL PHARMACOLOGY

2.0 PHASE I

2.1 Objectives: To determine human tolerance of a new agent and, when technically feasible, its absorption, distribution, metabolism, and excretion. Early studies in patients to determine whether the drug exerts a pharmacologic effect on the disease may be considered late Phase I studies.

2.2 Subjects: Adults; age 21 and over. Women of childbearing potential, children, and individuals with serious disease usually should be excluded. May be confined "normals" or institutionalized patients. Should be informed volunteers and have normal baseline physical examinations, laboratory, and other clinical studies. Generally should require no concomitant medication.

If normal volunteers are used, it must be recognized that the findings may have little relevance when compared with some psychiatric patients, particularly in their ability to tolerate side effects or larger doses of the drug.

2.3 Setting: Usually a confined setting in which provisions are available for close supervision and treatment on a 24-hour basis.

2.4 Investigator: An especially competent individual with experience in clinical pharmacology, psychiatry, or internal medicine who will conscientiously carry out frequent and thorough evaluation of all study subjects.

2.5 Study Design: Prior to receiving the study drug, all subjects should receive no other drugs for a period appropriate to insure that there will be no metabolic or symptomatic carryover effects. For example, if the previous drug has a relatively short half-life,

the necessary drug-free period may be only several days while, for phenothiazines, it should be much longer.

2.51 Single-Dose Study: It is beyond the scope of these guidelines to list the many variations in study designs that may be appropriate for the first trials of a new drug in humans. However, each should encompass the following general principles: First dose in the first subject should be minimal and based on animal toxicity studies (e.g., 1/5 to 1/10 of the maximum nontoxic dose in animals); monitoring before and after each dose (hence, term single-dose study); determination that each dose level is safe before administering a higher single dose to same or other subjects; interval before repeating same or higher dose in same subject sufficient to insure "washout" (when feasible as determined by animal and evolving human pharmacokinetic and pharmacologic data).

The following, which is to serve only as an example of one possible study design, may be accomplished with approximately six subjects in an open trial. The first subject receives a minimal single dose determined as stated above. The same subject is then given increments at three-day intervals. The second subject receives the maximum dose given the first subject. Each additional subject then starts at the maximum dose administered to the prior subject. All subjects continue to receive increments of the drug to the maximum tolerated dosage. Monitoring is conducted of clinical, laboratory, and when feasible, pharmacokinetic parameters at baseline, prior to receiving increments, and at follow-up intervals.

2.52 Multiple-Dose Study: Preferably double-blind with placebo control. Numerically balanced or imbalanced parallel group design may be used, e.g., 20 drug vs. 20 placebo, 12 drug vs. 8 placebo, or 30 drug vs. 10 placebo. Dosage increased at appropriate intervals (if possible, based on pharmacokinetics of single-dose study). As a precaution, may use subgroups starting at day or week intervals. When treatment emergent symptoms or side effects prohibit further increases, cut back to maximum tolerated dose and continue for at least 14 days. By this method, a total duration of approximately six weeks is usual, although it may often take longer. In any case, the duration of administration should be appropriate to the intended duration of further studies with recognition that additional and longer term safety data will evolve from them.

2.6 Assessment: Safety and/or toxicity usually assesed by: baseline and repeated extensive physical examinations; vital sign assessments; laboratory tests to assess the hematopoietic, liver, renal, and cardiovascular system. It is beyond the scope of guidelines to list all the specific tests that might be indicated. The type, extent, and frequency of testing will depend in part upon the type of drug, the preclinical information available, evolving information, and the eventual intended use of the compound. However, follow-up data usually should be obtained at least weekly during administration and whenever possible for at least one week after discontinuation of drug.

2.7 Documentation: Basically all study documentation should include the following:

a. Subject identification, demographic data, pertinent medical history, and vital statistics.

b. Pre- and post-treatment physical examination results.

c. Details of administration of medication and of dosage adjustment.

d. All behavioral and emotional effects observed or reported.

e. All adverse reactions reported or observed, the date, the severity and duration of such reactions, the investigator's judgment of whether drug-related, the control measures utilized, and the results of such measures.

f. All laboratory reports, including normal ranges for particular lab used.

2.8 Absorption and Metabolism: When initial safety studies have been completed, the absorption and half-life of the drug should be determined, if feasible. Such data might be relevant to the evaluation of safety and efficacy. More complicated metabolic studies can be postponed until Phase II when the utility of the drug will be more certain.

3.0 PHASE II

3.1 Objectives: In Phase II the overall objectives are:

3.11 To identify conditions or symptoms which may be therapeutically responsive to the drug.

3.12 To estimate the appropriate clinical dosage and duration of effect.

3.13 To identify adverse effects

3.2 Sample Selection: Early Phase II studies can be conducted with in-patient or out-patient adult males (and females who are not of childbearing potential), depending upon the degree of clinical supervision possible. Patients should be continually monitored for safety of the drug, as described in the Phase I outline.

As determined by comprehensive clinical and laboratory evaluations, patients evaluated early in this phase should require no concomitant medication and have no organic disease that may obscure clinical observations, laboratory tests, or interpretations.

The term "anxiety" has many different meanings. In the context of evaluating the therapeutic effects of antianxiety agents, "anxiety" refers to states of manifest anxiety. These states are characterized by the following:

Subjective experiences:

1. Feeling nervous, jittery, jumpy
2. Feeling fearful, apprehensive, anxious, panicky
3. Fears of fainting, screaming, losing control, crowds, places, disaster, death
4. Avoiding certain places, things, or activities because of fear
5. Feeling tense or keyed up

Muscular or motor phenomena:

6. Tense, aching muscles
7. Trembling, shaking
8. Restlessness, fidgeting

Autonomic phenomena

9. Heart beating fast or pounding; chest pain

10. Trouble catching breath, air hunger, smothering, lump in throat, choking

11. Sweating, especially armpits, palms, soles of feet

12. Cold, clammy hands

13. Dry mouth

14. Dizziness, faintness, lightheadedness, weakness

15. Tingling feelings in hands or feet

16. Stomach "gas", nausea, upset stomach

17. Frequency or urgency of bladder or bowels

Patients selected for trials of antianxiety agents should show the first two manifestations plus at least three others in the list above. A sufficient quantity of the pathology, including "moderate" or higher levels on the first two manifestations, should be present to warrant treatment and to allow room for improvement. This should be documented by initial scores which exceed specified levels on one or more of the quantitative criteria of drug effect. Although the specification of minimum levels at this time has to be largely arbitrary, mean levels in normal and pathological samples are available for some measures and can provide guidance (see "Assessment of Change").

Manifest anxiety as described above can occur in association with a variety of other symptoms and in a variety of diagnostic conditions. Anxiety may respond differently to medication in different clinical contexts. Two reasonable strategies are open to the investigator in coping with this problem. First, he may deal with anxiety strictly as a diagnostic entity, i.e., anxiety neurosis, and select his sample accordingly. Second, he may deal with manifest anxiety as an affective state occurring under many clinical circumstances and select his sample accordingly. The strategy that the investigator adopts will define the population to which he can generalize his results.

The second strategy should be limited to a diverse group of primarily psychoneurotic patients. Present knowledge suggests that anxiety may respond differently in patients with psychosis, borderline states, severe behavior disorders, addictions, or serious somatic or psychosomatic disorders. Antianxiety agents, therefore, should be studied separately in these conditions.

The mixture of anxiety and depression, even in psychoneurotic patients, also presents a special problem. Although the two states frequently occur together in varying degrees, most patients can be classified as primarily anxious or primarily depressed, both on clinical interview and on quantitative measures that assess the two effects. Claims that an agent is effective for both anxiety and depression should be supported by results in primarily anxious and in primarily depressed patients, rather than in some unspecified mixed states.

The response of manifest anxiety to medications depends not only on the accompanying pathology but may also depend on a variety of other circumstances including:

1. Duration of symptoms

2. Age of first onset

3. Number of prior episodes

4. Previous treatment and response to it

5. Prior medication

6. Type of subject: "normal," out-patient, or in-patient

7. Type of clinical practice

8. Age

9. Sex

10. Indications of social class, such as education and occupation

11. Race

The investigator should describe his sample systematically in terms of the above characteristics. He also may limit his sample to patients with specific characteristics, especially in small controlled Phase II studies, where heterogeneity presents the greatest problems. As determined by comprehensive clinical and laboratory evaluations, patients evaluated in late Phse I and early in Phase II should require no concomitant medication and have no organic diseases that may obscure clinical observations, laboratory tests, or interpretations.

3.3 Setting: As noted above, the setting within which anxious patients are treated may affect their response to antianxiety agents. Certain institutional settings, e.g., prisons, should be used only when there is assurance that the setting does not significantly affect the response to the drug.

Psychiatric hospitalization may be contraindicated for neurotic patients. Evaluation of antianxiety drugs in in-patients can be misleading. Therefore, initial out-patient studies may be done by investigators with extensive experience, facilities, observational and follow-up capabilities which assure patient safety. Each trial situation must assure provisions for the safety, care and adequate evaluation of study subjects as related to available and evolving information concerning the study drug.

Anxiety is often self-limiting and episodic and can respond to a variety of non-drug interventions. Response can be affected by such factors as: concomitant treatments including psychotherapy, physician's attitude and involvement with the patient, and environmental fluctuations. These factors should be kept as similar as possible between treatment groups.

3.4 Investigators: The investigators should be experienced in evaluating psychiatric drugs and in the conduct of clinical trials. They should have ready access to the appropriate population group for whom the drug may be indicated.

3.5 Design: Patients should be assigned to treatments at random. When safe and feasible, each study subject should have a drug-free period for several days prior to receiving the study medication. The number of days would depend upon the type and duration of any prior medication which the subject had been receiving. If the patient may be habituated to the prior medication, ample time should be allowed for any withdrawal effects to subside spontaneously. If the patient's symptomatology after washout has dropped below the selection criteria, the patient should be dropped.

Initially, several uncontrolled trials may be desirable to allow investigators sufficient flexibility to explore all possible aspects of a new drug's activity and to allow for the determination of an appropriate dosage range for use in double-blind studies. The open studies may be of small sample size. However, it may be desirable for clinical measures and selection of patient sample to be consistent between investigators in order to facilitate the interpretation of results.

It must be kept in mind that information obtained from these early open studies can only form hypotheses. Hypotheses evolved from open studies can be confirmed or refuted only by controlled double-blind studies. In at least some studies, the investigational new drug should be compared to a matching placebo control to establish its efficacy. Other studies may include only an active treatment control or both.

Packaging and coding of medications should be performed on an individual patient basis rather than on a treatment group basis. Intake of the study medication should be monitored at each visit by careful inquiry and by counting medication that is returned. Other psychoactive drugs are to be avoided. If other drugs are used, this should be carefully documented.

Parallel groups, crossover, intensive, and other designs may be used. The planning of this and other aspects of these studies should, whenever feasible, involve extensive consultation with a biostatistician. There should be full awareness of the advantages, disadvantages, and criteria for validity regarding each possible design before selecting one. Comparisons between two active drugs usually require much larger sample sizes than active drug-placebo comparisons.

There can be no requirements stated regarding specific studies or observations to determine a drug's potential for abuse since methodology at this time is still in the developmental stage. Nevertheless, even though problems of interpretation are recognized any indications of tolerance, such as a need for increasing doses to maintain effectiveness, should be reported. Observations of patients after discontinuation of the study drug may also yield useful information. It should be kept in mind that apparent signs of withdrawal may be due to reemergence of symptoms related to the basic disorder.

3.6 Dosage

3.61 Exploratory, open studies: After selection of initial dosage based on all previous data (including pharmacokinetic), dosage in open trials is usually increased until a satisfactory therapeutic response is observed. If adverse effects are a significant problem, further increases may be precluded and dosage reduction or discontinuation is indicated.

3.62 Double-blind, controlled studies: Dosage may be fixed. However, because of individual metabolism and tolerance, a flexible dosage may be necessary. A specified range may be used (as determined in earlier trials) within which adjustments are made individually according to specified clinical criteria.

The mode of administration, dosage range, schedule of administration, and criteria for dosage adjustment should be stated in each protocol.

3.7 Duration: The duration of individual clinical studies may vary from days to weeks depending on their purpose. The basic activity of antianxiety agents usually can be established in trials of two to four weeks' duration. However, to allow for a further assessment of safety beyond that obtained in Phase I, at least one of the first several Phase II studies should be continued for six weeks with appropriate laboratory monitoring if preliminary data have indicated satisfactory support of efficacy.

3.8 Assessment of Change: Baseline observations should be carried out in all patients immediately before their initiation into the study. The frequency of follow-up evaluations may then vary from days to weeks. However, since anxiety may remit rapidly under effective treatment, the second evaluation generally should take place within a week. The rater should not have access to previous ratings when he makes later ratings.

The selection of criteria for evaluating treatment response is crucial. Pertinent areas for assessment include the patient's subjective experience of symptoms; his functional status, social adjustment, or role performance (worker, spouse, parent, etc.); his physiological status; and his performance on a variety of specific tasks.

Possible observers of the patient's condition include the patient himself, family members or friends, the treating clinician, other professionals participating in the patient's care, such as nurses and independent research raters. A patient, if at all possible, should be rated by the same person in the same setting throughout his course in the study.

Specific classes of criteria that have been found useful in evaluating antianxiety agents include:

Global rating scales

Global rating scales can be used by various observers to summarize the patient's condition. The condition to be rated and the time period covered should be clearly specified. The ends of the scale should be well defined. The number of points offering maximum discrimination is not well established.

Two types of global rating scales are in common use. One refers to the patient's present status (absolute), often on a seven-point scale ranging from no pathology to most severe pathology, as measured against the rater's entire previous clinical experience.

The other global scale refers to the patient's change since the start of treatment or since the last visit (relative), often on a seven-point scale ranging from marked worsening (-3) through no change (0) to marked improvement (+3).

Symptom rating scales

A number of standardized rating instruments are available for patients, clinicians, or other observers to record the intensity of many symptoms associated with psychoneurotic states, including anxiety and depression. Some of these scales assess several dimensions in a common metric.

Target symptom ratings

These measures provide a tailormade instrument for the individual patient: The chief symptoms presented by the patient at the beginning of treatment are rated for their intensity initially and through the course of treatment. The key complaints can be selected by patient or clinician, and ratings can be made by the corresponding observer. Although these ratings have high relevance to the individual patient, group analyses encounter the problem of mixing dissimilar symptoms.

Mood adjective checklists

These measures provide quantitative assessments of various mood dimensions, including tension/anxiety and depression. Usually they are rated by the patient, although they can be used by other observers as well.

Other measures

Indirect verbal measures: These measures provide estimates of affective states and other areas of function through spoken or written verbal samples from the patient analyzed for content by carefully trained technicians according to procedures specified in detail. Although these procedures are less subject to some of the distortions encountered in direct measures, they are difficult and time-consuming.

Ratings of social adjustment: These measures cover such areas as work performance, interpersonal relations, and function as a spouse and parent. Some are designed primarily for use by a relative, but the clinician or the patient himself may rate others.

Indirect clinical measures: These measures include observations of the clinician's medication guesses, patients' reports of favorable and unfavorable life events, and rates of premature termination and their reasons. Interpretation of such measures tends to encounter problems of clinical relevance.

Physiological measures: Some of these measures address themselves primarily to the level of sympathetic arousal and include heart rate, respiratory rate, skin resistance or potential, forearm blood flow, and palmar sweating. Others, such as EEG patterns, which tap CNS activity directly, also have been employed to identify drug effects. Although these measures offer greater objectivity, interpretive problems remain with regard to clinical relevance.

Anxiety states by definition are primarily subjective. Functional disturbances associated with anxiety usually are less marked than in other conditions. The relationship between clinical anxiety levels and more objective physiological and behavioral measures is not well established. The patient generally takes the initiative in seeking, maintaining, and terminating treatment. Treatment contacts generally are on an ambulatory basis and limited to a single clinician. On account of these circumstances, the most important and practical measures of anxiety states focus on symptoms as evaluated by patient and treating clinician.

Criteria of drug effects in clinical trials of antianxiety agents should include at least a global rating and a standardized symptom rating by the patient. It is desirable for additional measures to be included also.

In selecting specific measures, the following properties of each measure should be considered:

Reliability

Reliability refers to the consistency of measurement. Measures that remain very stable over long periods of time are likely to tap characterological traits rather than affective states amenable to medication. Consistency on repeated occasions over brief time periods and among multiple observers on the same occasion are clearly desirable.

Validity

Validity refers to the ability of the measure to reflect the property of interest, in the present case, anxiety. The most practical test of a measure's validity in this situation is its demonstrated ability to discriminate between the effects of established antianxiety agents.

Comprehensiveness of content

It is useful for a measure to provide estimates of clinical variables besides anxiety, especially depression and psychotic trends, in order to define quantitatively the context within which anxiety is being studied, to document exclusion criteria, and to reveal other potential drug effects. Although separate instruments can serve these purposes, a common metric has an advantage.

Normative data

Normative data on patient and nonpatient samples aid the investigator in (a) setting minimum levels of anxiety for including patients in his study, (b) setting maximum levels of other pathology acceptable for including patients in the study, and (c) interpreting the effectiveness of treatment in reducing pathology.

Ease of administration

Measures requiring less time and effort on the part of the observer, particularly the skilled professional, have an advantage. Measures designed for patient self-report are more effective when comprehension requires minimal education. This involves both conceptual and linguistic level. More specific items are more readily rated by all observers. Scales with fewer points are easier to rate. Simplicity of scoring is an advantage.

In addition to the ratings recommended above, provisions should be made to record systematically in all treatment groups the emergence of new symptoms which may represent adverse side effects. The evaluation and documentation of new symptoms should include their severity, duration, likelihood of association with the study medication, and action taken in response.

Physical examinations and clinical laboratory tests also should be included. These should be basically the same as those for Phase I but may be modified in accordance with Phase I findings.

3.9 Interpretation: Several excellent discussions are available on the topic of interpreting the results of clinical drug trials. These papers document the many serious problems involved and raise doubt that any specifiable procedure can take account of all the possible contingencies. The following suggestions with respect to specific problems of interpretation are offered tentatively:

Criteria of drug effect

Drug trials usually include a variety of criteria, which may respond differently to the drug. Efficacy should be judged in terms of the most central criteria, and these should be specified in advance. They should include the clinician's global rating and the clinician's and the patient's ratings on the subjective, motor, and autonomic phenomena of anxiety. Efficacy should be claimed only in terms of those areas of psychopathology that consistently demonstrate a drug effect.

Statistically nonsignificant drug-placebo differences or similar mean responses to drug and placebo in a properly designed and executed study should raise doubt about the efficacy of the drug in that study unless the observed placebo response was unusually strong. Significant drug-placebo differences on criteria not specified in advance should be regarded as heuristic observations that require confirmation in additional studies.

Significant differences on quantitative criteria should be demonstrable between the (adjusted) final levels or the regressions of final on initial levels in the drug and the placebo groups. Significant differences also should be demonstrable in the number of individuals showing improvement, no change, or worsening in the two treatment groups.

4.12 Long-term safety studies: To establish the safety of a new antianxiety agent when given daily for 3 to 6 months or longer with particular regard to the nature, incidence, and control of adverse side effects.

4.2 Sample Selection: The same comments apply as in Phase II. In Phase III, larger and more heterogeneous samples may be studied. In such studies, the investigator may employ sample characterisitics in his analysis of results to provide more precise estimates of drug effects, to learn more about the effects of the sample characterisitics themselves, and to delineate the most responsive subsamples.

After controlled studies have indicated the drug's basic antianxiety activity, later studies in this phase may be established to explore its efficacy and safety in patients with organic disease requiring concomitant medication.

Females of child-bearing potential may be included if results of animal reproductive and teratologic studies are satisfactory.

Children may be studied in separate trials provided the clinical picture warrants drug administration and safety is assured to the fullest extent possible. Geriatric samples, generally, also should be studied separately, as responsiveness in different age groups can be very variable.

The issue of overriding importance is this: Claims of effectiveness should be confined to the conditions and samples actually studied and should be supported by well-specified documentation.

4.3 Setting: May be many, e.g., in-patient, out-patient, private practice, psychiatric or nonpsychiatric, and specialty, or G.P.

4.4 Investigators: Because the extension of claims into other areas involves a variety of types of investigators, it is important to consider the investigator's capability and experience in evaluating and working with patients with anxiety, his provisions for needed safety precautions, his access to laboratory facilities with suitable controls, and the appropriateness of his clinical setting to allow for valid drug evaluation.

4.5 Design: Of primary importance during Phase III of the drug's evaluation are controlled studies designed to fully confirm the drug's basic antianxiety efficacy. The design guidelines are generally the same as those discussed in Phase II. However, adjustments may be made in controls, duration of study, dosage, and design which do not interfere with validity (biostatistical consultation recommended) to accomodate greater variations in purpose of studies, settings, investigators, and subjects as discussed under the respective headings in Phase II. For example, patients who respond to an initial week of placebo medication may be excluded, and the criterion measures at that point may be taken as baseline readings in analyses of change over the subsequent period when patients have been randomly assigned to drug and placebo.

When it is concluded that the drug's basic antianxiety efficacy has been clearly established by controlled studies, consideration may be given to undertaking further studies on an open trial basis. These, of course, carry with them the inherent risk, due to lack of a control group for comparison, of encountering difficulties in interpretation of unexpected findings. However, such findings, as stated previously, can lead to forming hypotheses which must be confirmed or refuted by reviewing already completed studies or establishing further controlled studies. In providing further experience with the drug (often under conditions of usual medical practice), these studies can be important in providing corroborative support of efficacy demonstrated by well-controlled studies and in adding valuable data regarding safety of the new drug. This is particularly true when a number of investigators working independently obtain similar findings.

Sample size

Significant drug-placebo differences should be demonstrable in studies with 30-50 patients per medication group, i.e., in a sample size likely to be encountered in the individual practitioner's caseload. Groups of less than 20 patients per treatment are most unlikely to show significant drug-placebo differences.

There is no objection to combining data from several studies with similar protocols and forms to increase sample size, if necessary. (In many instances this procedure entails increases in heterogeneity which more or less offset the gains in numbers.) A collaborative study, however, does not obviate the need for replicating results.

Number of studies

Five well-designed and executed controlled studies of 30-50 patients per medication group should suffice to demonstrate efficacy. When the patients selected are reasonably homogeneous as to clinical picture, the significance level obtained by combining the significance probabilities of the drug-placebo differences from the five studies should be at least five percent, and all studies should show mean drug-placebo differences in the expected direction.

Comparisons with standard drug

The case for efficacy of a new antianxiety agent is strengthened if it exceeds placebo and also equals or exceeds a standard medication, such as chlordiazepoxide or diazepam, in effectiveness in two studies each employing three treatment groups.

Amount of improvement

The case for clinically significant efficacy is further strengthened if the amount of improvement with the new medication is substantial. Any of the following observations on the major measures of pathology would provide supportive evidence: (1) The proportion of patients improved with the drug divided by the proportion improved with placebo is at least 1.5; (2) The drug reduces pathology to a level that does not differ significantly from the general population norm; (3) The drug-placebo difference accounts for at least 5 percent of the total variation in outcome.

Deviations from protocol

All deviations from protocol, including patients who terminate, should be explained fully when results are submitted. Specific patients to be excluded from the analysis of results should be identified before the medication code is broken.

Analyses should be employed that maximize the number of subjects included. For example, one set of analyses may include all patients who met admission criteria and returned for at least one post-treatment visit.

4.0 PHASE III

4.1 Objectives

4.11 Therapeutic studies: Extension of comparative controlled studies to fully confirm in heterogeneous patient populations the drug's basic antianxiety activity and to provide more specific information about symptoms and the patient types in which the drug is especially effective.

Patients also may be evaluated in studies related to other than the basic antianxiety claim. For each of these other areas, a few double-blind studies should be sufficient. Prior to these, it may be necessary to carry out several open studies to familiarize the investigator with the drug's activity and appropriate dosage range in a particular therapeutic situation or special population. In these populations (such as the geriatric), carefully edited clinical pharmacology and therapeutic trials (i.e., they need not include all the variables of other studies) may be carried out to establish tolerance, efficacy, and safety.

Long-term safety studies may be on an open trial basis or may be of a controlled parallel group design. For controlled studies imbalanced groups may be used, e.g., 30 drug and 20 controls.

Data regarding long-term safety may also be obtained from a number of studies rather than from a single formally structured one. For example, provision may be made for patients in therapeutic trials to continue on the drug if it is indicated.

4.6 Dosage: In addition to dose considerations mentioned in Phase II, geriatric patients deserve special attention and consideration since antianxiety drugs are often extremely useful in the management of these patients. Older patients may not tolerate these drugs as well as younger adults if the same dosage is given. It is, therefore, necessary to establish in dose-response studies the dosage that gives adequate clinical response with a minimum of adverse side effects.

In long-term safety studies dosage should be of at least the level expected for eventual general therapeutic use. Allowance for adjustments to individual tolerance, of course, must be provided.

4.7 Duration: The duration of therapeutic trials (Phase III) may vary as in Phase II and, particularly regarding open trials, often may be longer. After baseline determinations in all cases, follow-up evaluations usually become less frequent as evolving data permit. The duration of long-term safety studies is usually 3 to 6 months. Baseline observations should be carried out in all cases, and follow-up evaluations are usually at least monthly.

4.8 Assessment: This is generally similar to that outlined for Phase II. Valid adjustments are permissible as indicated.

Some special examinations may be needed in the geriatric age group, e.g., EKG's or blood sugars may be needed since many of the patients will have concomitant organic disease.

4.9 Special Studies: The special studies considered in this section cannot be designated as requirements since the necessity for these or others will vary according to such factors as the pharmacologic and toxicologic background of a new agent, evolving data from clinical studies, and the drug's eventual intended use. Studies of a new drug's effects on psychomotor (and other) performance, on physiological correlates of anxiety, and on content analysis of verbal behavior may be carried out. Ongoing research in this area, hopefully, will significantly contribute to the more objective study of anxiety.

Methods of operant conditioning with the observation and rating of relevant behaviors may be worthy of consideration.

There may be value in studying the effects of antianxiety agents in non-patient subjects with anxiety. Such "models" of anxiety can be found among "symptomatic volunteers" (distressed persons who are not currently in treatment but tend to have histories of chronic disturbance with repeated, unsuccessful, prior efforts at treatment) or among "normal volunteers" with higher than usual levels of anxiety, particularly in stressful situations.

Long-term studies of safety and efficacy also present the opportunity to look for possible tolerance and dependence. These issues present special problems of design and interpretation.

It is beyond the scope of these guidelines to consider all the possible special studies that might be indicated concerning specific chemical compounds.

REFERENCES

1. Taber, ben Zion. Proving New Drugs: A Guide to Clinical Trials. Geron-X, Inc., Los Altos, Calif., 1969.

2. Harris, E.L., ed. The Principles and Practice of clinical Trials. Edinburgh and London: E & S Livingstone, 1970.

3. Levine, Jerome; Schiele, Burtrum; and Bouthilet, Lorraine, eds. Principles and Problems in Establishing the Efficacy of Psychotropic Agents. U.S. Government Printing Office, Washington, D.C., PHS Publication No. 2138, 1971.

☆ U.S. GOVERNMENT PRINTING OFFICE: 1977— 730-433/1604