DDT Vol. 6, No. 23 December 2001

Novel treatments for anxiety and depression: hurdles in bringing them to the market

Raimund Buller and Valérie Legrand

Anxiety and depressive disorders are the most common psychiatric conditions. The medical need for newer, better-tolerated and more efficacious treatments remains high. However, drug development is time-consuming and has a high rate of failed or inconclusive trials. Improvements in study design, investigator training and early proof-of-concept studies are being discussed as means to decrease failure rates and the duration of development. So far, no uniformly applicable 'magic formulas' for success have been discovered. The most promising approach to overcome these hurdles appears to be a sound study design carried out by experienced professionals in the clinic and in industry.

Raimund Buller and Valérie Legrand
Quintiles France
3–5 Rue Maurice Ravel
F-92564 Levallois-Perret
Paris, France
*tel: +33 1 41 27 72 17
fax: +33 1 41 27 72 00
e-mail: raimund.buller@quintiles.com

▼ Depressive and anxiety disorders are the most frequent mental disorders. They exist together more often than as separate syndromes. More than 20% of the adult population suffer from these conditions at some time during their life. The impact on productivity and quality of life is significant because the disorders affect subjects during their most productive age periods. The World Health Organization (WHO) predicts that depression will become the second leading cause of premature death or disability worldwide by the year 2020. During recent decades, the overall risk of suffering from depression has increased and the age of onset has decreased. Despite numerous public campaigns, the lay audience and the medical community have been very slow in realizing these facts[1]. As shown by some recent European research, depression is still underdiagnosed and undertreated. In general, the majority of patients with major depression (59%) do not receive any antidepressant medication at all. Even when antidepressants are prescribed they are often given at inadequate doses or for insufficient periods of time[2]. Less information is available on the current status of diagnosis and treatment of anxiety disorders; however, many feel that the situation is comparable to that in depression, if not worse. Data from the USA indicate that only 30% of those with a probable depressive or anxiety disorder received appropriate treatment. Appropriate care was received by only 19% of those in primary care but by about 90% of those visiting mental health specialists[3]. The majority of patients are not diagnosed accurately but they receive outdated broad diagnostic labels, such as 'anxiety neurosis', which do not help to select the most efficacious treatments. Consequently, many patients with anxiety disorders take benzodiazepines (BZDs) for long periods although they are most effective in short-term use: long-term use of BZDs has also been shown to increase the risk of drug dependence[4]. A smaller but increasing group of anxious subjects is now treated with antidepressants, which have been shown to be effective in short-term and long-term treatment of several anxiety disorders[5]. Approximately two-thirds of the anxious or depressed patients respond to the currently available treatments but the magnitude of improvement is still disappointing. In depression, 50–85% relapse during the remission phase or develop a new depressive episode following recovery. Hence, merely 15% of depressed individuals remain continuously well after their first episode[6]. A prospective study to assess the frequency and course of mental disorders, conducted in a cohort of young, healthy Swiss men who were followed over >25 years (Zurich Study), found that the majority of depressed patients suffer from recurrent depressive episodes, but only 11% of the cases had a single episode and 13%

 1359-6446/01/$ – see front matter PII: S1359-6446(01)02043-8

developed a chronic depression[7]. Depression is, therefore, an illness that often requires prolonged treatment. Prognosis for the more severe anxiety disorders appears to be comparable to that in depression and many patients require long-term treatment for their anxiety.

Current and future treatments for anxiety and depression

Major anxiety disorders include generalized anxiety disorder (GAD), panic disorder (PD), agoraphobia, social phobia (now also called 'social anxiety disorder' – SAD), post-traumatic stress disorder (PTSD), obsessive compulsive disorder (OCD) and adjustment disorders. For many years, the BZDs have been the mainstay of treatment for anxiety disorders. Their efficacy is best established in acute anxiety, which is found mainly in adjustment disorders with anxious mood, GAD, PTSD, panic disorder and some phobias. Long-term efficacy of BZD has not been established in these conditions. Benzodiazepines appear to lack efficacy in OCDs (Ref. 8), although a small study with clonazepam suggests that subjects with OCD could benefit from treatment with this compound[9]. Although Alprazolam has shown some efficacy in patients with milder depression[10], the efficacy of BZDs in the treatment of more severe forms of major depression has not been established. However, they are frequently used concomitantly in depressed patients to treat associated symptoms of anxiety or insomnia. Benzodiazepines, or other drugs that act through the GABAergic system, will remain the drugs of choice for acute anxiolysis, in other words, for acute anxiety states[11,12]. In the treatment of anxiety disorders they are now slowly replaced by antidepressants as most of the so-called 'antidepressants' are not only efficacious in depression but also in the acute and long-term treatment of GAD, OCD, SAD, PD and PTSD. Therefore, the market for antidepressants has tremendous growth potential and could even double in size during the next 10 years, creating revenue expectations of > US$14 billion in the USA by the year 2006 (data according to analysts Frost & Sullivan; http://healthcare.frost.com). Although the patent expiry of Prozac has led to a rapid replacement by generic fluoxetine and an 80% drop in market share for the Lilly drug (Indianapolis, IN, USA), industry analysts predict no major impact on the pricing of other still patent-protected drugs. Many treated patients and their prescribers will stay loyal to their brand-name compounds and will resist being switched from other selective serotonin re-uptake inhibitors (SSRIs) to the generic fluoxetine simply because of cost reasons. However, generic versions of antidepressants could help to increase the number of patients starting treatment and could, therefore, lead to a higher sales volume for antidepressants during the coming years.

First- and second-generation antidepressants

The effects of first-generation antidepressants, such as monoamine oxidase inhibitors (MAOI) and tricyclic antidepressants (TCA), were discovered by a combination of serendipity and circumspect clinical observation. All currently marketed antidepressants are believed to act principally on neurotransmission at the level of the synapse, by either blocking the re-uptake of monoamines, by inhibiting transmitter degradation or by binding directly to specific receptors[13]. With better understanding of these mechanisms, drugs with more specific profiles were developed, such as the group of SSRIs (e.g. fluoxetine, fluvoxamine, sertraline, paroxetine and citalopram) and, more recently, drugs that specifically inhibit norepinephrine re-uptake[14], such as reboxetine[15]. Other, newer compounds have mixed effects on several neurotransmitters (e.g. venlafaxine[16] and milnacipran[17]) or simultaneously act on receptor subpopulations as antagonists (α-2 antagonist mirtazepine[18,19] and 5-HT_2 antagonist nefazodone[20]). Concurrently, safer reversible inhibitors of monoamine oxidase A (RIMA), such as moclobemide, have also been developed, although they have a minor role in the treatment of depression. In general, the prescription of modern antidepressants is easy because many compounds permit once- or twice-daily dosing, nonetheless, it is questionable if newer drugs are more efficacious than the older ones. Meta-analyses show that the treatment response in depression has not dramatically changed and that most antidepressants appear to produce a 50% or greater reduction of depression scores on common rating scales in approximately 50–60% of patients. The SSRIs as a class have not been found to be more efficacious than TCAs, which, by contrast, could have better efficacy in hospitalized patients[21]. Although some factors are known to have an impact on outcome (co-morbidity, severity, chronicity and type of symptoms), it has not been possible to predict individually who will eventually benefit from treatment and to select the right drug for the right patient. However, compared with previous generations of antidepressants, newer compounds are much less toxic in overdose and have better safety and tolerability profiles. It is still an open question whether more recently introduced drugs, such as venlafaxine and mirtazepine, have a faster onset of action or greater efficacy, whereas at the same time methodological questions on how to best determine early onset of action are still being debated[22–24].

Future drugs

Attempts to improve the efficacy of antidepressant treatments have focussed on reducing the latency of the response and on finding strategies for treatment-resistant

cases. Several studies or individual reports have described the effects of augmentation with pindolol[25] and combinations of antidepressants with dopaminergic agents[26,27] (e.g. atypical neuroleptics). This has created expectations that greater efficacy in subgroups could be achieved with different mechanisms of action and newer treatment strategies are, therefore, directed at pharmacological targets, which go beyond the monoamines. Because of this, drug research in depression and anxiety is currently focussing on various neuropeptide receptor ligands[28,29]. Several compounds are being developed that target endocrine abnormalities in depression, such as the overactivity of the hypothalamic–pituitary–adrenal (HPA) axis[30]. Clinical studies are on the way to test the effects of corticotropin-releasing hormone (CRH) antagonists. One compound has shown some preliminary evidence of efficacy[31] but, in the meantime, has been discontinued for reasons related to the safety of the drug. Therefore, whether the CRH antagonist approach holds promise for patients with depression or anxiety is still an open question. Other targets include neuropeptide Y (NPY) and substance P, and their G-protein-coupled NPY- or neurokinin (NK)-receptors[32,33]. Substance P and/or NK-1 antagonists have been investigated as analgesics for several years and some of them are being tested in depression and anxiety disorders, for example, social phobia[34,35]. Recently, one NK-1 antagonist (MK-869) has shown efficacy in a study in moderate to severe depression[36], although this finding has not been replicated in a second trial because of high placebo response rates[37]. However, the drug appears to have been remarkably well-tolerated. By contrast, drugs that act on cholecystokinin (CCK) receptors, which also belong to the family of peptide receptors, have not been effective in the treatment of clinical anxiety despite positive data from animal experiments[38]. Other drug targets could be found in receptor subpopulations (like the serotonin 5-HT_2 receptor family)[39]. Future possibilities of better modelling of G-protein receptors might be particularly helpful for drug design not only in psychopharmacology[40]. Among drugs that target 5-HT_2 receptor subpopulations are compounds such as deramciclane[41], which is currently in development for the treatment of anxiety. Antidepressants for the new millennium[42] could also be drugs that interact with *N*-methyl-D-aspartate (NMDA) receptors[43,44]. Finally, future treatments might try to bypass the classical receptors altogether to interact directly with secondary messengers and sub-receptor components of the transmitter systems (e.g. the linkage between receptors and G proteins). Antidepressants have already been shown to modify gene expression and future targets might be found among intracellular transcription factors, such as CREB (cAMP response element binding protein), which appear to be involved in the production of nerve growth factors (NGF), which might be necessary to preserve neuronal integrity[45].

More specific drugs but for smaller patient groups?

Although there are some studies suggesting that treatments based on more specific targets could ultimately lead to improved efficacy, the debate whether this is, in fact, the case will continue for some time. Depression as it is currently conceptualized is a heterogeneous condition, which is clinically defined by descriptive criteria and not by pathogenesis or aetiology. The outcome is variable and is influenced by the interplay of psychological, environmental and genetic factors. More specific treatments might actually lead to a better response in subpopulations but could result in less favourable overall outcome in the total population of depressed subjects. Therefore, in the future, depressive disorders might have to be reclassified into subgroups that are defined by biological variables, before more specific treatments can be applied. However, although postulated for many years, no such biological classification of depression is currently in sight. This is not surprising, because the underlying pathology of depression is still not sufficiently well understood.

The role of genomics?

High expectations are currently linked to the future role of genomics, which is hoped to lead to the next therapeutic revolution and has, therefore, attracted major investments from drug companies. Even in the most promising case, over the next 5–8 years genomics will not have a major impact on the development of drugs for anxiety and depression. Whether genomics will actually have such a revolutionary effect in multiple-gene disorders, such as anxiety or depression with all the possible modifications of the illness caused by environmental or psychological factors, appears questionable. Current antidepressants are basically so easy to use because they have excellent tolerability and because they only treat the symptoms of depression regardless of their aetiology. This is what makes them so appropriate for the treatment of large populations of sufferers. Marketed antidepressants are not attempting to correct or cure the specific underlying causes of the illness. This might not be possible anyhow, as there are probably variable and manifold causes and interactions between 'causal' factors, which would be difficult to target with just a few compounds. Because of their broad activity (and, thus, limited specificity) antidepressants are also successfully used to treat the symptoms of several anxiety disorders. They even have a place in the treatment of conditions such as chronic pain or fibromyalgia[46]. In fact, many

compounds already have obtained licences for indications other than depression. Because virtually millions of patients can be treated with an antidepressant (although not every single subject will respond) the drugs have now become 'blockbusters' with potential peak sales well beyond US$1 billion per year.

By contrast, highly specific and customized future treatments must aim at smaller subgroups of only several thousand depressed patients. It is hoped that they will be more effective than current drugs and that this would justify higher prices. Although the development costs and risks for these newer compounds would probably still remain high they could generate much lower revenue than the current 'broad spectrum' drugs with their admittedly limited efficacy. There have been speculations that a therapeutic approach based on specific (genetic) abnormalities could actually prove to be quite devastating to the current pharmaceutical industry as 'common diseases would no longer be common'[47]. This could well be the end of the blockbuster drugs and the beginning of a radical change in the drug development process, possibly also requiring changes in regulations or duration of patent life.

Current problems in drug development for anxiety and depression

Because of the limited efficacy of current drugs, the search for better compounds has to proceed. Without specific chemical or biological tests, the diagnosis of depression continues to be based on classical descriptive, rather than aetiological, criteria. Furthermore, over the coming years, general practitioners and private practice psychiatrists who prefer easy-to-use drugs with broad activity will remain the major care providers for anxious and depressed subjects. For these reasons, development programmes for new anti-anxiety and antidepressant drugs will follow conventional tracks for quite some time. Drug developers will have to face the same risks and complications that were so characteristic of all past clinical developments in depression and anxiety[48,49]. Treatment effects, although clinically relevant, are often modest and, therefore, require large samples to be detected with statistical significance. Placebo response can be high, thus making demonstration of efficacy difficult[50]. As a result, definite proof-of-efficacy is not possible in early phases of development but only in relatively large Phase III trials. Still, up to 50% of these studies are inconclusive and do not show any statistically significant differences between test drug, placebo and active comparator. Failure of a compound at this stage, or the need to repeat studies, is a very costly complication.

Costs are also driven by the regulatory requirements[51]. The authorities are mainly concerned with safety and the possibility of false-positive results, therefore, they ask for at least two positive studies as proof of acute efficacy. Justification of the proposed dose and demonstration of long-term efficacy for anxiety and depression (at least for a registration in the European Union) are also necessary. For the safety databases a minimum of 1500 patients have to be exposed to the test compound and even larger datasets could be needed when the benefit is small or only experienced by a fraction of the patients treated. Long-term safety should be demonstrated in 300–600 patients treated for six months and in at least 100 patients treated for one year. Consequently, clinical development programmes usually include 3000–5000 subjects, making the development a risky, time consuming and expensive exercise.

Overcoming the challenges in antidepressant development

The pharmaceutical industry is still facing rising development costs, whereas the time spent for clinical development has not been reduced significantly during recent years. By contrast, innovative compounds with novel mechanisms of action nowadays rapidly lose their exclusive status because competitor drugs with similar mechanisms of action are introduced into the market almost simultaneously. These factors severely threaten the profitability of new compounds. Several approaches have been considered to eliminate drug candidates with questionable therapeutic benefits as early as possible, or to minimize development time, costs and risks. Animal models and proof-of-principle studies could help identify weak candidate molecules in early phases. More refined clinical trial methodology and rigorous reliability training for investigators have been suggested to increase the likelihood of success in later phases of clinical development.

Animal models

Normal and pathological anxiety in humans is a complex phenomenon involving physiological, behavioural, cognitive and psychological changes. Application of animal models to human anxiety might not be consistent across the range of diagnostic categories because they are conceptionalized in psychiatric classifications, such as the Diagnostic and Statistical Manual of Mental Disorders (DSM-IV; Ref. 52). Most models can only mimic behavioural aspects (like avoidance), which are also seen as part of some of these disorders. Animal models for anxiety are based on either unconditioned responses, conditioned responses, 'fear' states induced by drugs, brain stimulation or adverse social stimuli, but not all appear to be equivalent in terms of the elicited emotional state. Models using unconditioned responses are variations of the theme of

locomotion in novel environments (like the elevated maze). The theory behind these models is that fear or anxiety mediates inhibition in behaviour because anxiolytics can reduce inhibition. Models based on conditioned responses (like the Geller–Seifter conflict test[53]) are more complicated but allow better control over the animal's behaviour. Several of the older models, which could correctly identify BZD-like drugs, have failed to detect drugs with different mechanisms of action. The best recent example is buspiron[54], whose anxiolytic potential was missed by conventional screening. Newer models, such as those based on social interactions, have since been developed to predict correctly the anxiolytic activity of drugs with serotonergic or other non-GABAergic mechanisms.

Like in the case of anxiety, animal models for depression suffer from the fundamental problem that there is no convincing evidence for the existence of these disorders, even among the most developed non-human primates. The essential affective and cognitive characteristics of depression, such as guilt, worthlessness, pessimism and suicidal ideation, are unlikely to be found in animals. Therefore, some animal models focus on observable behaviours, which resemble those seen in depressed patients. Reward or punishment, stress or anatomical lesions can manipulate these animal behaviours. Current depression models, at best, have some predictive validity (i.e. response to antidepressants, although often with acute, instead of chronic, treatment) and face validity (i.e. similarity in behaviours, especially in anhedonia models, with a decreased response to pleasurable stimuli). However, they lack construct validity (i.e. similarity of underlying causes for human depression and changes in animal behaviour). According to Leonard[55], an animal model of depression should show behavioural changes that simulate those occurring in depressed patients; symptoms should respond (normalize) when an antidepressant is administered; because antidepressants only work when given as chronic treatments, the animal model should respond optimally only with chronic administration. Most animal models do not meet all of these requirements. However, they can still be useful for identifying agents with effects on neurotransmission, especially in the serotonergic and noradrenergic system. Not surprisingly, these models might be unable to identify clinically effective antidepressants with different mechanisms of action. By definition, rats have less developed brains than humans and their behaviour is less complex. In the development of new antidepressants it might be possible to investigate other physiological changes, which are also seen in depressed subjects (like those in endocrine, immune or neurotransmitter systems) in appropriate rodent models of depression. There is some evidence that lesions in the amygdala are responsible for eliciting some of these changes in the rat[56], which could also have clinical implications. Animal models have contributed to the development of newer antidepressants and to the understanding of their biological activity and will certainly remain important in the screening of new drug candidates. However, because of their limitations in predicting precisely if and how effective a compound will be, they are not safeguards against failures in clinical development.

Proof-of-principle studies in anxiety and depression

Benzodiazepines were developed as drugs for the (acute) treatment of anxiety, irrespective of its aetiology or pathophysiology. As such, they have therapeutic value in organic anxiety associated with medical disorders (such as myocardial infarction), in acute stress reactions and in anxious schizophrenic or depressed patients. With their broad spectrum of activity, BZD-like compounds have also been tested successfully for their anxiolytic activity in normal volunteers in situations evoking anxiety or mental stress (like anxiety in anticipation of minor surgery or dental treatments). Currently, it is not possible to develop drugs as 'general anxiolytics' or for the treatment of broad indications, such as 'anxiety'. Rather, drugs are licensed only for the treatment of specific anxiety disorders as defined by accepted classification systems. Attempts to find experimental models for these conditions in healthy human volunteers have, so far, not been successful. Even drug-induced 'panic-like' anxiety states and panic-attacks observed in patients appear to be different in their symptomatology or response to treatment. Furthermore, therapeutic activity in clinical anxiety disorders is often only observed with continued treatment for several weeks. Similarly, for depression there are currently no study designs to find definite evidence of whether a drug will become an effective antidepressant in patients during earlier phases of development by way of so-called proof-of-principle or proof-of-concept studies. Previous hopes that sleep electroencephalography (EEG) and hormone challenges could identify antidepressants within a short time period in just a few subjects and during an early phase of development have proven to be premature. Brain-imaging techniques also currently have little to offer in terms of treatment outcome or diagnosis of depression (other than identifying organic causes of depression). Receptor-binding studies and pharmaco-EEG could help to select dose ranges for a trial but do not qualify as proof-of-principle. At times, data from self-rating scales given to healthy volunteers in Phase I pharmacology studies, or depression ratings in studies with non-depressed patients, are thought to demonstrate mood-enhancing properties of a drug. Such

findings have to be taken with extreme caution: as with the effects observed in patients suffering from anxiety disorders, mood changes in depressed patients usually occur only after 1–2 weeks of treatment; clinical pharmacology studies rarely last longer than 1–7 days. Most clinical depression scales are validated for depressed samples, but not necessarily for normal subjects or for patients with other illnesses. Finally, short-term mood improvement, or even euphoria, does not predict antidepressant action as can be seen from the experience with BZDs and, especially, with drugs of abuse (amphetamines, cocaine), which are not effective as treatments for depression.

Therefore, definite proof-of-efficacy requires testing new compounds in patient populations; however, because of the modest effect sizes commonly seen with treatment, studies with larger samples are necessary. These studies are usually only conducted late in Phase II or even in Phase III trials. Treatment duration is normally 6–8 weeks in depression and up to 12 weeks in SAD and OCD. Frequently, investigators and ethics committees request that treatment for responders be extended up to six months. This requires significant investments to generate the necessary preclinical chronic safety data and to produce and formulate the relatively large quantities of the test article. Therefore, discontinuation of a programme at such a late stage because of lack of efficacy is a rather costly failure.

Faster drug development by biotech and virtual drug companies?

The drug development process in the pharmaceutical industry has been criticized as being too slow and bureaucratic. There have also been complaints that large pharmaceutical companies lack innovation because creative individuals are leaving the industry and are being replaced by functionaries, and research is driven by lawyers, financial experts, salesmen and market strategists who are 'completely unable to develop new ideas'. Talented, creative minds are working more often in smaller companies and especially in the biotech industry. However, these companies frequently suffer from important financial restrictions and find it difficult to pay for full development and market introduction of an antianxiety or antidepressant drug. Even small to mid-size pharmaceutical companies have to look for partners to share the costs and risks of clinical development. To attract investments or interest from potential partners, rapid evidence is required to support the continued development of a compound. Thus, short-term objectives of finding financial support could absorb a lot of the creativity and get into conflict with the long-term targets of bringing a new drug to the market rapidly. Proof-of-concept studies would be an ideal tool to generate the necessary interest but for many indications these are not available. Most small companies also cannot afford to hire experienced, talented drug developers for every therapeutic area of research and, therefore, they often rely on outside consultants, who are certainly expert clinicians but not necessarily experts in drug development. Weak expert advice has resulted in poorly designed trials that are overloaded with assessments. In an attempt to find answers to a multitude of questions in the shortest possible time and in as few patients as possible, these trials risk being inconclusive as they regularly suffer from insufficient statistical power. When, for cost reasons, they also either lack a placebo group or an active control group, the interpretation of ambiguous results is nearly impossible. Furthermore, smaller companies often only have a limited number of compounds (or even only a single one) under development and their financial situation (including the stock price) depends on generating 'good news' for these compounds. This creates a significant risk for bias in the interpretation and presentation of results.

Collaborations

Optimally, collaboration between small and large companies should be a source of synergy and creativity. However, a successful collaboration requires partners with sufficient expertise on both sides willing to enter into an honest exchange of the available information. Collaborations can be difficult to organize; usually they have their own dynamics and depend on the right kind of 'people chemistry'. Whether, even under the best circumstances, partnerships can bring compounds to the market faster has not yet been proven. On the contrary, codevelopment scenarios can become quite risky, costly and time-consuming, especially when trials or complete parts of a programme need to be repeated for reasons of insufficient methodology.

Recently, so-called 'virtual development companies' or 'virtual teams' within large companies have entered the drug-development process. They are called virtual because they do not produce or market drugs but only manage parts of the clinical development. They also do not enter into clinical trial monitoring or marketing. Virtual companies usually in-licence compounds that have been dropped from the portfolio of traditional pharmaceutical companies for development in Phase I or II trials. This helps the originator companies to save resources but still allows them to retain the rights to their compounds, which can be bought back once they have demonstrated their potential to justify full development. If they fail to meet certain predetermined milestones the financial risks for the originator company are obviously minimal. The tasks of the virtual company are to design the clinical studies, to

organize them (most often through a contract research organization) and to find the financial means to fund them. The return on this investment is paid from fees generated by successfully developed compounds, which are sold back by the originator, or to a third party at a premium price. Although this model might be able to reduce some financial risks, it is probably not helpful in speeding up the development process. The managerial and financial negotiations and decisions involved in licensing-out the compound, or in buying it back, can be quite time-consuming. Virtual drug development potentially also suffers from conflicts between activities that serve short-term goals, such as the identification of suitable drug candidates, attracting investors, selling the drug back to the originator, and those that support the long-term goal of getting a new drug to the market. As time is precious, there is always a temptation to cut corners and, because of financial pressures, the interpretation of study results can also become biased. Although the experience with virtual development is still limited, it is questionable whether it will reduce development time or costs of a full development. The true benefits appear to be that drugs get a chance for development, which would otherwise not have made it into a portfolio. The model could also reduce some of the financial risks of a failed development for the original owners of compounds. However, the price for a successful compound must include a compensation for the cost of the failed candidates and, therefore, will reduce the profitability for its original owners.

Improved trial methodology

Significant efforts have gone into activities designed to decrease the number of failed studies and, thus, development time and costs. Improved trial methodology, in general, aims to reduce the inherent variability of the condition studied and protects drug-placebo differences so that they can be detected with adequate sample sizes. A short comparison between animal research and human trials should illustrate some major methodological differences. Animal studies need smaller samples (n = 10–20 per group) than trials in depressed or anxious patients (70–150 subjects per group) because of reduced variability in outcome. In animal experiments, test groups are usually homogeneous (inbred strains, with similar age and weight, and a defined gender ratio) and live in standardized conditions (e.g. cage, food and daylight cycles). Endpoints are few and well-defined objective measures (e.g. time spent with forced swimming, volume of sweetened water ingested). Treatments and assessments are followed with high precision. Patient groups in therapeutic trials are much more heterogeneous (e.g. differences in age, social and genetic background) and live in different types of environment (e.g. work, family, activity and diet). Endpoints in human research are multiple and complex (global severity, depression or anxiety ratings, disability scores, level of social impairment, quality of life and life satisfaction). Treatment and assessments are followed with less precision (non-compliance, drop-out and subjective ratings). Furthermore, treatment is only one of the factors that has an influence on outcome. Frequently, the treatment factor, including specific and nonspecific effects, does not account for much more than 10–15% of the variability of the response. The remaining variability is explained by other prognostic factors (e.g. demographic and social factors, chronicity and severity) and, to a significant extent, by chance.

Several common reasons for failed trials or false-negative results are related to the characteristics of the patients entered. If the level of severity in the trial differs significantly from the anticipated level, either only a few subjects (severity too high), or almost everyone (level too low), will respond to treatment. Treatment-resistant subjects or those with relevant co-morbid conditions might show reduced drug response. Non-compliance, underdosing and the use of concomitant medications or other concurrent treatments and inadequate washout from previous medications could also reduce treatment differences. Therefore, protocols often require a minimal level of severity for inclusion (e.g. a score of at least 18 or 21 points on the Hamilton Depression Rating Scale). Excluded are patients with co-morbid conditions (e.g. psychotic disorders, abuse and dependence, anxiety disorders) and the use of other psychotropic drugs. Design-related reasons for false-negative data are inadequate sample size, insufficient treatment duration, wrong dosage and inappropriate clinical endpoints. Important investigator-related reasons for inconclusive studies could be non-compliance with the protocol, biased selection of subjects, lack of precision in diagnosis and unreliable or invalid assessment of outcome. To reduce investigator biases in patient selection for trials, and to increase the recruitment rates, a pre-screening based on patient self-ratings via telephone using an interactive voice response system (IVRS) has been tried recently. Potentially eligible subjects were referred to the nearest trial centre to be entered into the study. Although the recruitment rates were much faster, no psychiatric study using this methodology has, so far, been able to demonstrate a treatment effect. Whether there is a real advantage in this methodology or if it is just a way of going ‘faster to disaster’, as some have said, is not clear at the moment.

Problems with the precision in assessments could also stem from inadequate rating instruments. The primary outcome measures (pivotal scales) need to have good

psychometric quality (validity, reliability, the sensitivity to detect change and the ability to discriminate between active drugs and placebo). For international trials, valid and reliable translations must be available. Rating instruments should be acceptable for the scientific community and regulatory authorities. Many of the scales that are currently used (e.g. Hamilton Anxiety Scale, Hamilton Depression Rating Scale[57], Montgomery–Åsberg Depression Rating Scale[58]) were developed long before the current classification systems came into use and are, therefore, only partially able to measure such new constructs as GAD, major depression or dysthymia. Many of the scale items lack specificity and their total score is a problematic measure of severity as it is unclear if the items truly tap only one factor or illness dimension. However, these instruments are still widely used and have even attained the questionable status of a gold standard, and will probably not be replaced in the near future. For some of the anxiety disorders, newer and more specific scales, like the Liebowitz Social Anxiety Scale[59] (LSAS) or the Yale–Brown Obsessive Compulsive Scale[60] (Y–BOCS) have been developed. To overcome the weaknesses of the instruments, additional refinements have been proposed, such as the use of standardized structured interviews and operational criteria or anchors for scoring. All these are intended to reduce information and interpretation variance. Extensive guidelines and rater trainings should help to reduce the 'threshold' variance (i.e. differences among assessors in rating the presence or severity of a symptom) and the influence of errors.

Methodological improvements are necessary but obviously not sufficient. By contrast, there are currently also some important trends, which could significantly slow down the development of new antianxiety or antidepressant drugs or even completely alter their development process. More often we hear that there appear to be enough equipotent antidepressants with good safety characteristics and with low rates of side effects. Benkert and coworkers[61] even suggest that superior efficacy versus a standard drug in acute and continuation treatment of moderate and severe depression should be required for the approval of future antidepressants. Recently, the use of placebos as a control in depression trials has also come under heavy criticism. Based on the latest revision of the Helsinki Declaration[62], an increasing number of ethics committees and investigators now argue that placebo-controlled trials in anxiety or depression have become invariably unethical when a known effective treatment is available. Opponents to this rigorous position have pointed out that trials comparing only two active treatments (so called non-inferiority trials) would need to expose much larger groups of subjects to new drugs of still unproven benefit but with potentially dangerous side effects.

Temple and Ellenberg[63] have discussed in detail the scientific implications. They explain why regulatory authorities, such as the Food and Drug Administration (FDA), in the case of antidepressants, still require evidence from placebo-controlled trials because anxiolytics and antidepressants belong to a group of drugs (such as analgesics and antihistamines) with significant assay sensitivity problems, in other words, a reduced ability of studies to distinguish between active and inactive treatments. In the case of antianxiety or antidepressant drug trials, ~30–50% of modern trials do not distinguish a known drug – used in the studies as active comparator – from a placebo. Some of the reasons for these failures have already been listed and include factors such as protocol violations, use of concomitant treatments and spontaneous improvement. However, even with all the current knowledge, it is still not possible to specify a particular study population, treatment protocol or sample size that will regularly avoid these problems and always identify an active agent. Although all study protocols appear to provide unambiguous inclusion and exclusion criteria, well-accepted outcome parameters, adequate sample sizes and clearly defined rules for conducting the trial, they have not reduced the likelihood of producing a failed or inclusive study. Therefore, for the moment it is wise to include at least four or five potentially pivotal trials as a safeguard to come up with the two required positive studies proving the efficacy of a compound.

The awareness factor – better understanding for clinical research

Under these circumstances, top management in several pharmaceutical companies has become sceptical as to whether the development of drugs for the treatment of anxiety or depression is still a wise investment with good chances of success. Although methodology might have only a limited influence on success, it is not justifiable to completely turn against methods. Specific adequate methodology will remain an important factor but needs to be supplemented by a human awareness factor. Clearly, it is necessary to select valid and reliable rating scales and to have sufficient data on their psychometric properties. However, the actual performance of the instruments in a given study with a given patient will depend on the way it is handled by the investigator[64]: is there enough time for the assessment?; is the investigator careful and experienced? It is essential to have a scientifically sound research protocol, but the protocol also needs to be followed properly by investigators and study staff. The reason why drug research sometimes does not have the reputation of being part of serious science is that it is often not performed with the required meticulous efforts. Instead, some view clinical

studies and the significant fees involved merely as an easy way to get financial support and, therefore, believe that it might be permissible to cut corners and to take the easy road without doing any harm.

Good clinical practice (GCP) guidelines have provided a strong basis for ethical and scientifically sound research and have coherently summarized the responsibilities of all parties involved in clinical research. The main aim of GCP is to protect subjects and to assure that the results are credible. Following these guidelines guarantees global acceptance of data and makes the results extremely valuable. However, many investigators do not yet see GCP as an opportunity to increase the importance of their individual contributions to clinical research and rather, they find GCP to be a rigid bureaucratic nuisance in their daily activities. This attitude could be a result of, in part, the fact that GCP guidelines were the results of an effort involving pharmaceutical industry and regulatory authorities, whereas investigators remained on the sidelines and did not participate directly (except for a few physicians working in the pharmaceutical industry who had previously been investigators). Since the earliest versions, GCP guidelines have remained external to researchers. Consequently, as can be observed in many investigator meetings, the interest in GCP training and the active knowledge about the GCP guidelines is usually quite low. Clear understanding or, even better, a sense of ownership, are needed not just a ritualistic training. The way GCP is followed (or not) is usually an excellent indicator for the quality and the dedication with which a site carries out research. At the moment, and for the near future, the demand for research sites by far exceeds the number of high-quality sites. Therefore, it is necessary, and possible, to increase the quality in many centres. The most important (human) knowledge factor that determines performance is the investigator's own value system and attitude. If investigators believe in providing high-quality data they will usually make a realistic estimate on recruitment rates and ensure that they spend the time and resources to meet this target. Centres with good performance are in an excellent position to select the most interesting trials and to determine in which ones they want to participate. If sponsors want to have access to the best sites they in return have to offer high quality, for example, in terms of study design, study conduct, project management, monitoring and medical or scientific support.

A new model for clinical research

Focussing on the awareness factor really means that all persons involved (i.e. project manager and research assistants in the sponsor company; investigators, nurses and trial co-ordinators at the site) have a clear understanding of the problems that are intrinsic to clinical research and that they strive to overcome these problems together. Investigators often do not understand that their role in a trial is quite different from what they do in routine treatment outside a clinical trial. They might have an excellent clinical understanding of depression but that does not mean that they are also experts in depression trials. Usually, significantly more time is spent with a study patient than in clinical routine to provide information to the patient for obtaining consent and to collect the data required by the study protocol. In many countries, a diagnostic classification, like the DSM-IV, is rarely used other than for scientific purposes, and rating scales or structured interviews are not a common part of routine care. Hence, investigators often lack practice and tend to underestimate the time and efforts needed to use these instruments properly. Furthermore, during the interaction with the patient in normal practice physicians apply all available means to improve the subject's condition. Although it can be extremely rewarding for the physician to 'cure' every patient, it also means that the ability of a study to detect efficacy decreases dramatically. Therefore, in a clinical trial, investigators are asked to limit their therapeutic activities to minimize the influence of nonspecific factors (e.g. support, encouragement, understanding and explanations), which normally contribute to the placebo response. Researchers should realize that it might be more appropriate for the study to withdraw a patient when the response is insufficient, rather than to maximize nonspecific treatment factors to force a response. It should also be sufficiently clear to investigators that ambiguous or inconclusive results are much more problematic from an ethical, scientific and financial point of view, than studies that are clearly negative and, therefore, aid in the decision to terminate the development of an inefficacious drug. Many investigators also still seem to underestimate the importance of reporting safety parameters. Especially in the area of anxiety and depression it is not rare to see under-reporting of adverse events. Similarly, vital signs are measured without paying enough attention to the necessary circumstances (e.g. blood pressure only taken when seated but not after rising). However, safety data are just as relevant as efficacy data because they are the basis for a risk-benefit assessment of a new compound and, therefore, are vital for decisions concerning further development and, ultimately, registration of a drug. Drugs with better safety features have contributed significantly to progress in the treatment of depression. Finally, insufficient observation and incomplete reporting of safety data is potentially very dangerous for the trial subjects and future users of the drug. It is also a clear violation of the investigator responsibilities as described in the GCP guidelines.

Conclusion

The development of drugs for the treatment of anxiety disorders and depression is currently seen as difficult, if not uncontrollable. However, participants need to realize that the overall success of the work can only improve when all involved guarantee the high quality of their contribution. Fast development of antianxiety and antidepressant drugs is basically a scientifically and methodologically sound development that requires clear understanding of the problems for all concerned and their ability and willingness to overcome them together. There is no magic formula for the quality of a programme; it rather depends on a continuous effort and is determined on a day-to-day basis; weaknesses in one part cannot be compensated by the other participants. A bad design cannot be corrected by the meticulous work of a dedicated investigator. By contrast, the damage done by centres working in a manner that is less than perfect cannot be counterbalanced by scales of the highest reliability. Every trial needs to be taken seriously and must have the appropriate type of centres and the right study team led by a dedicated project manager and guided by experienced scientific and medical advisors or principle investigators.

References

1 Hirschfeld, R.M. *et al.* (1997) The national depressive and manic-depressive association consensus statement on the undertreatment of depression. *J. Am. Med. Assoc.* 277, 333–340

2 Lépine, J-P. *et al.*, on behalf of the DEPRES Steering Committee (1997) Depression in the community: the first pan-European study DEPRES (Depression Research in European Society). *Int. Clin. Psychopharmacol.* 12, 19–29

3 Young, A.S. *et al.* (2001) The quality of care for depressive and anxiety disorders in the United States. *Arch. Gen. Psychiatry* 58, 55–61

4 Uhlenhuth, E.H. *et al.* (1999) International study of expert judgment on therapeutic use of benzodiazepines and other psychotherapeutic medications: IV. Therapeutic dose dependence and abuse liability of benzodiazepines in the long-term treatment of anxiety disorders. *J. Clin. Psychopharmacol.* 19 (6 Suppl. 2), 23S–29S

5 Skaer, T.L. *et al.* (2000) Anxiety disorders in the USA, 1990 to 1997. *Clin. Drug Invest.* 20, 237–244

6 Thase, M.E. (1992) Long-term treatments of recurrent depressive disorders. *J. Clin. Psychiatry* 53 (Suppl.), 32–44

7 Angst, J. *et al.* (1996) Recovery from depression: risk or reality? *Acta Psychiatr. Scand.* 93, 413–419

8 Montgomery, S.A. (1994) Pharmacological treatment of obsessive compulsive disorders. In *Current Insights in Obsessive Compulsive Disorder* (Hollander, E. *et al.*, eds), pp. 215–225, John Wiley & Sons

9 Hewlett, W. *et al.* (1992) Clomipramine, clonazepam and clonidine treatment of obsessive compulsive disorder. *J. Clin. Psychopharmacol.* 12, 420–430

10 Laakmann, G. *et al.* (1995) Treatment of depressive outpatients with lorazepam, alprazolam, amytriptyline and placebo. *Psychopharmacology* 120, 109–115

11 Uhlenhuth, E.H. *et al.* (1995) International study of expert judgement on therapeutic use of benzodiazepines and other psychotherapeutic medications: II. Pharmacotherapy of anxiety disorders. *J. Affect. Disord.* 35, 153–162

12 Uhlenhuth, E.H. *et al.* (1999) International study of expert judgement on therapeutic use of benzodiazepines and other psychotherapeutic medications: VI. Trends in recommendations for the pharmacotherapy of anxiety disorders, 1992–1997. *Depress. Anxiety* 9, 107–116

13 Vetulani, J. and Nalepa, I. (2000) Antidepressants: past, present and future. *Eur. J. Pharmacol.* 405, 351–363

14 Kent, J.M. (2000) SNaRIs, NaSSas, and NaRIs: new agents for the treatment of depression. *Lancet* 355, 911–918

15 Kasper, S. *et al.* (2000) Reboxetine: the first selective noradrenaline re-uptake inhibitor. *Expert Opin. Pharmacother.* 1, 771–782

16 Mendlewicz, J. (1995) Pharmacologic profile and efficacy of venlafaxine. *Int. Clin. Psychopharmacol.* 10 (Suppl. 2), 5–13

17 Spencer, C.M. and Wilde, M.I. (1998) Milnacipran. A review of its use in depression. *Drugs* 56, 405–427

18 Fawcett, J. and Barkin, R.L. (1998) Review of the results from clinical studies on the efficacy, safety and tolerability of mirtazapine for the treatment of patients with major depression. *J. Affect. Disord.* 51, 267–285

19 Thompson, C. (1999) Mirtazapine versus selective serotonin reuptake inhibitors. *J. Clin. Psychiatry* 60 (Suppl. 17), 18–22

20 Montgomery, S.A. (1996) Efficacy of nefazedone in the treatment of depression. *J. Psychopharmacol.* 10 (Suppl. 1), 5–10

21 Anderson, I.M. (2000) Selective serotonin reuptake inhibitors versus tricyclic antidepressants: a meta-analysis of efficacy and tolerability. *J. Affect. Disord.* 58, 19–36

22 Nierenberg, A.A. (2001) Do some antidepressants work faster than others? *J. Clin. Psychiatry* 62 (Suppl. 15), 22–25

23 Blier, P. (2001) Pharmacology of rapid-onset antidepressant treatment strategies. *J. Clin. Psychiatry* 62 (Suppl. 15), 12–17

24 Stahl, S.M. *et al.* (2001) Evidence of early onset of antidepressant effect in randomized controlled trials. *J. Clin. Psychiatry* 62 (Suppl. 4), 17–23

25 Olver, J.S. *et al.* (2000) Pindolol augmentation of antidepressants: a review and rationale. *Aust. N. Z. J. Psychiatry* 34, 71–79

26 Ostroff, R.B. and Nelson, J.C. (1999) Risperidone augmentation of selective serotonin reuptake inhibitors in major depression. *J. Clin. Psychiatry* 60, 256–259

27 Shelton, R.C. *et al.* (2001) A novel augmentation strategy for treatment resistant major depression. *Am. J. Psychiatry* 158, 131–134

28 Griebel, G. (1999) Is there a future for neuropeptide receptor ligands in the treatment of anxiety disorders? *Pharmacol. Ther.* 82, 1–61

29 Hökfelt, T. *et al.* (2000) Neuropeptides – an overview. *Neuropharmacology* 39, 1337–1356

30 Holsboer, F. (1999) The rationale for corticotropin-releasing hormone receptor (CRH-R) antagonists to treat depression and anxiety. *J. Psychiatr. Res.* 33, 181–214

31 Zobel, A.W. *et al.* (2000) Effects of the high-affinity corticotropin-releasing hormone receptor 1 antagonist R121919 in major depression: the first 20 patients treated. *J. Psychiatr. Res.* 34, 171–181

32 Saria, A. (1999) The tachykinin NK_1 receptor in the brain: pharmacology and putative functions. *Eur. J. Pharmacol.* 375, 51–60

33 Stout, S.C. *et al.* (2001) Neurokinin$_1$ receptor antagonists as potential antidepressants. *Annu. Rev. Pharmacol. Toxicol.* 41, 877–906

34 Rupniak, N.M. and Kramer, M.S. (1999) Discovery of the antidepressant and anti-emetic efficacy of substance P receptor (NK_1) antagonists. *Trends Pharmacol. Sci.* 20, 485–490

35 Vassout, A. *et al.* (2000) NKP608: a selective NK-1 receptor antagonist with anxiolytic-like effects in the social interaction and social exploration test in rats. *Regul. Pept.* 96, 7–16

36 Kramer, M.S. *et al.* (1998) Distinct mechanism for antidepressant activity by the blockade of central substance P receptors. *Science* 281, 1640–1645

37 Enserink, M. (1999) Can the placebo be the cure? *Science* 284, 238–240

38 Argyropoulos, S.V. and Nutt, D.J. (2000) Peptide receptors as targets for anxiolytic drugs. In *Anxiolytics* (Briley, M. and Nutt, D., eds), pp. 151–175, Birkhaeuser

39 Roth, B.L. *et al.* (1998) 5-hydroxytryptamine$_2$-family receptors (hydroxytryptamine$_{2A}$, hydroxytryptamine$_{2B}$, hydroxytryptamine$_{2C}$): where structure meets function. *Pharmacol. Ther.* 79, 231–257

40 Flower, D.R. (1999) Modelling G-protein-coupled receptors for drug design. *Biochim. Biophys. Acta* 1422, 207–234

41 Kanerva, H. *et al.* (1999) Brain 5-HT2A receptor occupancy of deramciclane in humans after single dose administration – a positron emission tomography study. *Psychopharmacology* 145, 76–81

42 Skolnick, P. (1999) Antidepressants for the new millennium. *Eur. J. Pharmacol.* 375, 31–40

43 Wiley, J.L. (1997) Behavioral pharmacology of *N*-methyl-D-aspartate antagonists: implications for the study and pharmacotherapy of anxiety and schizophrenia. *Exp. Clin. Psychopharmacol.* 5, 365–374

44 Petrie, R.X. *et al.* (2000) The *N*-methyl-D-aspartate receptor, synaptic plasticity and depressive disorder. *Pharmacol. Ther.* 87, 11–25

45 Maubach, K.A. *et al.* (1999) Novel strategies for pharmacotherapy of depression. *Curr. Opin. Chem. Biol.* 3, 481–488

46 Schatzberg, A.F. (2000) New indications for antidepressants. *J. Clin. Psychiatry* 61 (Suppl. 11), 9–17

47 Horrobin, D.F. (2000) Innovation in the pharmaceutical industry. *J. R. Soc. Med.* 93, 341–345

48 Klerman, G.L. *et al.* (1994) Evaluating drug treatments of depressive disorders. In *Clinical Evaluation of Psychotropic Drugs: Principles and Guidelines* (Prien, R.F. and Robinson, D.S., eds), pp. 281–325, Raven Press

49 Bourin, M. (2000) Clinical methodology for testing of anxiolytic drugs. *Thérapie* 55, 147–153

50 Thase, M.E. (1999) How should efficacy be evaluated in randomized clinical trials of treatments for depression? *J. Clin. Psychiatry* 60 (Suppl. 4), 23–31

51 European Commission (1998) The rules governing medicinal products in the European Union. Guidelines, Vol. 3 (http://www.eudra.org/emea.html)

52 American Psychiatric Association (1994) Diagnostic and Statistical Manual of Mental Disorders, (4th edn), American Psychiatric Association

53 Geller, I. and Seifter, J. (1960) The effects of meprobamate, barbiturates, D-amphetamine and promazine on experimentally induced conflict in the rat. *Psychopharmacologia* 9, 482–492

54 Shayegan, D.K. and Stahl, S.M. (2000) Buspiron. In *Anxiolytics* (Briley, M. and Nutt, D., eds), pp 13–25, Birkhaeuser

55 Leonard, B.E. (1998) Animal models of depression. In *Antidepressant Therapy at the Dawn of the Third Millennium* (Briley, M. and Montgomery, S., eds), pp. 87–109, Martin Dunitz

56 Kelly, J.P. *et al.* (1997) The olfactory bulbectomized rat as a model of depression: an update. *Pharmacol. Ther.* 74, 299–316

57 Hamilton, M. (1960) A rating scale for depression. *J. Neurol. Neurosurg. Psychiat.* 23, 56–62

58 Montgomery, S.A. and Asberg, M. (1979) A new depression scale designed to be sensitive to change. *Brit. J. Psychiat.* 134, 382–389

59 Liebowitz, M.R. (1987) Social phobia. *Mod. Probl. Pharmacopsychiatry* 22, 141–173

60 Goodman, W. *et al.* (1989) The Yale-Brown Obsessive Compulsive Scale (Y-BOCS): Part I. Development, use and reliability. *Arch. Gen. Psychiatry* 46, 1006–1011

61 Benkert, O. *et al.* (1998) Existing therapies with newer antidepressants – their strengths and weaknesses. In *Antidepressant Therapy at the Dawn of the Third Millennium* (Briley, M. and Montgomery, S., eds), pp. 213–230, Martin Dunitz

62 World Medical Association (2000) Declaration of Helsinki: ethical principles for medical research involving human subjects. Edinburgh, Scotland, October 2000

63 Temple, R. and Ellenberg, S.S. (2000) Placebo-controlled trials and active-control trials in the evaluation of new treatments. *Ann. Intern. Med.* 133, 455–463

64 Bartko, J.J. (1991) Measurement and reliability: statistical thinking considerations. *Schizophr. Bull.* 17, 483–489