

DAVID HEALY

MANUFACTURING CONSENSUS

ABSTRACT. The Food and Drug Administration (FDA) has declared that it would be illegal to advertise as or in any way claim your drug to be superior to competitors on the market, which are up to 30 times cheaper. How does a pharmaceutical company market such a product? The answer is to enlist academics to form expert panels to construct guidelines and algorithms, or participate in Delphi panels and other exercises, which can be expected to prove that newer, more costly drugs produce cost savings. These academics do so on the basis of the existing clinical trial evidence—which supposedly the FDA has used to come to its verdict that the newer compound is no better than its competitors. However, where the FDA has seen the raw data, academics later see the published data. In between intervenes a medical writing exercise, which produces the first and most important piece of advertising for any pharmaceutical product—the randomized controlled trial infomercial. This paper explores how pharmaceutical companies manufacture an apparent academic consensus and, in so doing, gives a case study of the recent controversies surrounding the marketing of selective serotonin reuptake inhibitor (SSRI) drugs for adolescent depression.

KEY WORDS: adolescent depression, FDA, manufacturing consensus, pharmaceutical marketing, SSRI

BACKGROUND

Consider this excerpt from the 1993 Food and Drug Administration (FDA) medical review of Janssen Pharmaceutical's application to market the antipsychotic drug, risperidone (Risperdal):

We would consider any advertisement, promotion or labeling for Risperdal false, misleading or lacking fair balance under Section 502 of the Act if there is a presentation of data that conveys the impression that risperidone is superior to haloperidol or any other marketed antipsychotic drug product with regard to safety or effectiveness. [Mosholder 1993]

The early clinical trials undertaken by Janssen with Risperdal compared it to an older antipsychotic, haloperidol. In a similar way, all recently released antipsychotics, including olanzapine (Zyprexa), quetiapine (Seroquel), and ziprasidone (Geodon), were compared to haloperidol in key premarketing trials. All companies used a high dose of haloperidol that was no more efficacious than lower doses of haloperidol but did cause more side effects. The company rationale for using haloperidol was that haloperidol was supposedly the market-leading antipsychotic agent.

Culture, Medicine and Psychiatry **30**: 135–156, 2006.

© 2006 Springer Science+Business Media, Inc.

DOI: 10.1007/s11013-006-9013-3

Whatever the real rationale, it was generally accepted at the time these trials were conducted that newer agents stood their best possible chance of looking better in terms of key side effects if compared to the doses of haloperidol used in the trials.

In addition to gatekeeping the entry of a new drug to the marketplace, the role of a drug regulator is to regulate any claims a manufacturer might make as regards a new product in its advertising or detailing to doctors. This assessment by the FDA would appear to produce problems for any company that might wish to market Risperdal.

But regulators have no control over what key opinion-leading academics, in the case of Risperdal, professors of psychiatry, say in lectures or report in medical journals or elsewhere. The FDA, in addition, has no control over what assessments these academics might make in their roles as experts called on to contribute to an expert consensus on new versus older drugs. Shortly after Risperdal was launched, it was being widely touted by leading academic opinion leaders as being superior to older antipsychotics.

Aside from the perennial need to market the product, the 1990s brought a new hurdle for drug companies to vault. It was increasingly necessary to persuade clinicians and pharmacists that a new drug should be listed on hospital formularies that had been created to ensure that new agents would not be used without good evidence of cost-benefit returns. The formularies are notionally meant to be evidence based and cost-sensitive. A certain amount of trade-off was likely—if a new drug cost more but could show a real benefit over older agents, it would be included.

No convincing evidence has ever been forthcoming that any of the new “atypical” antipsychotics are superior to the older “typicals” in either safety or efficacy. A study completed in 2003 by the VA hospitals compared olanzapine (Zyprexa) and haloperidol in terms of both efficacy and tolerability and found no difference between them; olanzapine in this study, however, cost approximately 80 times as much as haloperidol (Rosenheck et al. 2003). Despite this lack of greater efficacy, olanzapine won a place on formularies after its launch in 1996 to the extent that it became the most profitable antipsychotic in the world.

In the absence of clear evidence from clinical trials sufficient to warrant claiming that a new drug is superior to an older drug, it would appear to be difficult to make the extra step to advocating that the newer agent is more effective to the point of warranting a potential 80-fold increase in expenditure. Nevertheless, shortly after their launch, Risperdal and other recently released antipsychotics were available on most hospital formularies in both the United States and Europe.

Pharmaceutical companies have clearly found methods of circumventing these difficult areas of marketing terrain. Circumvention is achieved by recruiting senior academics and institutions to their cause by means of three stratagems: consensus conferences, pharmacoeconomic modeling, and ghostwriting.

CONSENSUS CONFERENCES

Consensus conferences aimed at producing guidelines for clinical practice came into existence in the late 1980s (Sheldon and Smith 1993). A range of scientific bodies took up this development, which in the first instance appeared to be a means for academia/science to rein in the excesses of pharmaceutical company marketing departments. In psychiatry, groups such as the British Association of Psychopharmacology and the European College of Neuropsychopharmacology, for example, produced guidelines for the treatment of conditions from depression through schizophrenia. The involvement of such organizations may have been in part a shrewd effort to establish their political profile, but in a number of such organizations, the influence of key individuals with links to pharmaceutical companies is apparent.

At the same time, companies began to sponsor meetings aimed at producing expert consensus on issues such as the appropriate use of medication in schizophrenia. These company-sponsored meetings led to products that appear indistinguishable from non-company-sponsored guidelines or algorithms. While this might be thought of as an exercise designed to confound the recommendations of independent committees, in fact independent committees have come up with recommendations that barely differ from explicitly company-sponsored exercises.

Given the lack of evidence base for the superiority of the new antipsychotics, just how have all these guidelines ended up endorsing newer, more costly agents over older, less expensive, but equally effective ones? One such guideline system, the Texas Medication Algorithm Project (TMAP), offers some answers (Petersen 2004).¹

Risperdal was launched in 1994. The TMAP was instituted in 1995, initially funded by Janssen Pharmaceuticals (Johnson and Johnson), the makers of Risperdal. Soon afterward it attracted funding from all major pharmaceutical companies. The TMAP drew up a panel of consultants to produce an expert consensus on the use of antipsychotics and, later, on the use of antidepressants and mood stabilizers (Gilbert et al. 1998). Many of

these consultants had prior links to Janssen and the other major pharmaceutical companies operating in the mental health field.

The first set of TMAP guidelines concluded that the atypical antipsychotic medications Risperdal, Zyprexa, and Seroquel were the drugs of choice for the management of schizophrenia (Chiles et al. 1999). The second set concluded that newer patented antidepressants, such as Prozac, Paxil, and Zoloft, were the drugs of choice for the treatment of depression rather than older tricyclic antidepressants. Subsequently mood stabilizers such as Depakote and Lamictal have been endorsed over other treatments for bipolar disorder. In all these instances, the claims have been that the new drugs are safer, more effective, and better tolerated than the older agents. The expert panels then formulated a set of algorithms or care pathways for the treatment of schizophrenia, depression, and bipolar disorder based on these guidelines.

In a number of U.S. states, legislators have the powers to rule that algorithms and guidelines such as these must be applied in the care of any patients receiving treatment at public facilities. The logic here is that evidence-based guidelines and algorithms, if they really do reflect reality, can be expected to be cost-effective over time. The legislators faced with the question of adopting the algorithm and guideline proposals in Texas meet infrequently, are poorly paid, and are intensively lobbied. Not surprisingly perhaps, the TMAP was administratively endorsed in Texas, and as a result state hospital doctors were required to follow its algorithms and use these newer drugs first.

Researchers linked to the TMAP were also able to access the records of patients in state facilities, including prison hospitals and mental hospitals, and report on the cases that appeared to do favorably. These surveys produced data supporting the selection of Risperdal and Zyprexa, for instance, as first-line treatments for schizophrenia and, later, the selection of selective serotonin reuptake inhibitors (SSRIs) or other newer antidepressants over older treatments for depression. On this basis, the TMAP guidelines and algorithms began to be referred to as evidence-based guidelines and evidence-based best practices.

A related panel formulated a set of medication algorithms for children, which recommended new antipsychotics and antidepressants, such as Paxil (paroxetine), for the management of children's problems (Hughes 1999). In this case, not only was there a lack of evidence for the superiority of the newer over the older agents, but also there was essentially no evidence base for the recommendations other than a set of then unpublished clinical trials.

The TMAP algorithm and guidelines were subsequently marketed to other states and instituted by administrative decision in a number of cases.²

In this way very few people had effectively paved the way for the acceptance of these guidelines and algorithms in many states and produced a situation in which a growing cohort of patients treated in the public sector ends up being put on and maintained on these drugs. It will probably come as no surprise that within Janssen there was a special unit aimed at maximizing the effectiveness of the companies marketing in the public sector.

From TMAP to NICE

While the TMAP process appears to be close to egregious, something very similar happened in the socialized system of medicine in Britain. In the first place, opinion leaders in Britain were recruited to panels to produce evidence-based guidelines for antipsychotics. The experts invited to such meetings had no pressure put on them to come to a particular point of view. All of the publications of clinical trial data for antipsychotic drugs were made available to them, and they were encouraged to be evidence based.

Again, as with the TMAP, the results, despite the assessment of the FDA (which was unknown to the participating experts), must have been gratifying to the sponsoring company (Mortimer et al. 1998).³ The process involved no overt selling of named medications but, rather, a set of positions endorsing the use of antipsychotics in monotherapy regimens, at doses consistent with British National Formulary recommendations and in a manner that would avoid precipitating acute treatment-related side effects.

Subsequently, the National Institute of Clinical Excellence (NICE) was set up in Britain with a brief to make recommendations as to the most clinically effective and cost-effective treatments for both physical and mental illnesses. The NICE guidelines for psychiatric treatment are an essentially similar creation to the TMAP and earlier United Kingdom-based industry-sponsored guidelines: a consensus of expert views rather than evidence-based views. The process involves a small number of psychiatrists, psychologists, and other stakeholders in mental health such as psychiatric pharmacists collating evidence, preparing draft reports, and then sending these to selected experts for comments. Decisions are reached not by experiment or evidence but by agreement. The process takes into account prior algorithms and guidelines. And finally, as has been pointed out by the WHO, the NICE process operates within the constraints of company unwillingness to share the raw data arising from clinical trials (World Health Organization 2003).

The upshot of this in the case of the antipsychotics has been a set of guidelines indistinguishable from the ones drawn up by the TMAP or other company-sponsored guideline groups (National Institute of Clinical

Excellence [NICE] 2002). The NICE recommends the use of the new antipsychotics over the old, even though it acknowledges that it does so without having any evidence base for this. In fact, the NICE guidelines for schizophrenia flew in the face of then published evidence that new antipsychotics produce significantly higher death rates and significantly higher suicide rates than older antipsychotics or placebo, as well as a range of physical problems, from cardiovascular to endocrine disorders.

In a public health system such as the National Health Service, the NICE guidelines are implemented in a very similar way to the TMAP guidelines. The medical directors of hospitals will ordinarily seek to ensure that their clinical staff adhere to NICE guidelines, as one indicator that clinicians are pursuing the best possible clinical practice. As a direct result of the NICE, then, a much larger number of patients will end up being given new rather than old antipsychotics than would otherwise have been the case, with a probable resulting detriment in the collective patient health, brought about at vast cost. It is all but impossible for individual clinicians to opt out, as practicing outside the guidelines may not be regarded as evidence based.

The critical influence here lies with the clinical trials that supposedly form the basis for the guideline process. Newer agents almost invariably have more and larger trials than older agents, especially if this is for indications that have been “created” since the older drugs went off patent. Many older agents may have minimal trial data. Those constructing the guidelines rarely appear to take into consideration the fact that the larger the trials needed, the weaker the drug must be, and that, in general, trials are only needed when there are some doubts as to whether the drug actually works. But, even more critically, underlying data that might occasion a different set of conclusions may not be available in the published literature to those constructing the guidelines.

In fact an abundance of relevant data was publicly available in reviews published by the FDA for each of the new antipsychotics at the time of licensing, and in the case of suicides and deaths, many of the data were available in a paper on rates of suicides and suicidal acts in clinical trials with novel antipsychotics (Khan et al. 2001). These data show high rates of suicide on Risperdal and perhaps the highest rates of suicide in clinical trial history on Zyprexa (see Table 1). But the most surprising thing is that the paper by Khan et al. offers no figures for suicidal acts on Zyprexa, while it does offer figures for suicidal acts in the clinical trial programs for the other new antipsychotics. Against the background of possibly the highest suicide rates in clinical trial history, this absence of data on suicidal acts for Zyprexa is striking.

TABLE 1

Incidence of Suicides and Suicide Attempts in Antipsychotic Clinical Trials Drawn from FDA License Applications

	Number of Patients	Number of Suicides	Number of Suicide Attempts	All Suicidal Acts
Risperidone	2,607	9	43	2.00%
Comparator	621	1	5	1.00%
Placebo	195	0	1	0.50%
Olanzapine	2,500	12	?	?
Comparator	810	1	?	?
Placebo	236	0	?	?
Quetiapine	2,523	1	4	0.20%
Comparator	420	0	2	0.48%
Placebo	206	0	0	0.00%
Sertindole	2,194	5	20	1.14%
Comparator	632	0	2	0.32%
Placebo	290	0	1	0.34%
Ziprasidone	2,993	6	?	
Comparator	951	1	?	
Placebo	424	0	?	
Total				
New Antipsychotic	12,817	33		
Comparator	3,434	3		
Placebo	1,351	0		

Note. The data obtained by Khan et al. (2001) are supplemented here with data for suicidal acts on quetiapine and sertindole provided by the companies making them and on ziprasidone from FDA medical reviews. In contrast to Khan and coworkers' scrutiny of FDA medical reviews for olanzapine, my scrutiny suggests that the true figure for suicides on placebo in olanzapine trials was zero. Analyzing the data on suicides using an exact version Mantel Haenszel procedure and a one-sided test for significance yields a relative risk of suicide of 6.9 with a confidence interval of $(1.0825, \infty)$, $p = .03$, for suicides on new antipsychotics compared to suicides on placebo.

In the case of Zyprexa, the 2,500 patients noted in Table 1 were participants in five clinical trials. This problem with Zyprexa was not hidden by leaving trials unpublished. Quite the contrary. As of 2004, there had been more than 200 publications featuring these five trials, of which over 40 were significant publications, but none of these offer any indication that the suicide rate was so high or any indication of the suicidal act rate. For companies aiming to influence an expert consensus, unpublished trials are of little use.

PHARMACOECONOMICS

In the case of these newer agents, companies have also resorted to a set of pharmacoeconomic procedures. Pharmacoeconomics as a discipline began

in the 1970s, heavily subsidized by the pharmaceutical industry (Healy 1998). It basically involves estimating and comparing the costs of leaving a condition untreated against the costs of treatment. The original view of the first pharmacoeconomists was that the complications of establishing treatment effects and outcomes for psychotropic drugs across a range of domains of value in mental health meant that it would be impossible to apply the procedures of pharmacoeconomics to psychiatric conditions and treatments.

Nonetheless, the emergence of a set of SSRI antidepressants and atypical antipsychotics that could not be distinguished from older agents in terms of efficacy or tolerability, but that were associated with greatly increased costs, led to a flurry of pharmacoeconomic exercises. This is exemplified nicely by the emergence of supplements to major journals detailing a range of pharmacoeconomic approaches that probably did a good deal to smooth the marketing path of the SSRI antidepressants (Eccleston 1993).

One of these methods involved the establishment of Delphi panels of experts. Delphi panels invite experts to consider clinical trial data and estimate the likely translation from the actually published randomized trial evidence to possible outcomes in clinical practice if the drugs are adopted widely. These outcomes are then costed by economists working for the manufacturing company.

The participants in these exercises will again be unaware of assessments such as those made by the FDA or the data on suicide or death rates from trial programs. The invariable outcome of these proceedings has been sets of models indicating that treatment with newer agents costing 10 to 80 times more than older agents would in fact lead to savings in either for-profit health-care systems such as that in the United States or socialized medical systems such as the U.K. mental health system (Guest et al. 1996).⁴

No one seems prepared to say what the original exponents of pharmacoeconomics realized, namely, that short-term trials cannot be used for this purpose. This issue is now further complicated by something that would once have been all but inconceivable, namely, that in a growing number of cases critical aspects of the raw data are substantially at odds with the published data.

GHOSTWRITING

In the 1980s, pharmaceutical companies began to outsource functions such as the running of clinical trials and medical writing to other companies. Medical writing was outsourced to medical communication agencies. With this development, the practice of ghostwriting academic articles picked up

pace. Ghostwriting involves medical writers writing articles that subsequently appear under the apparent authorship of key opinion-leading academics who might have reviewed the piece before publication; ghosts traditionally are medical writers who receive no credit for their input. For some time it was believed that this form of medical communication was largely confined to journal supplements or peripheral journals (Healy 2003a, 2004). The first hints that the picture might be somewhat different came in the mid-1990s. Flanagin and colleagues, for example, reported in 1998 that up to 11 percent of articles published in six mainstream peer-reviewed journals involved the use of ghostwriters.

Recently a document became publicly available covering the coordination during the course of 1998 of medical articles on Pfizer's antidepressant Zoloft (sertraline) by a medical communications agency, Current Medical Directions (CMD). This has permitted the comparison of published articles written for Pfizer with other articles on Zoloft in terms of the impact factor of the journals in which they appeared, publication history of the respective authors, and subsequent citation rates of the respective series of articles.

The analysis showed that the journals in which Pfizer's articles were published had an impact factor three times greater than the journals in which other articles on Zoloft were published. The authors on Pfizer's articles had nearly three times more previously published articles cited in Medline and Embase than the authors of articles not linked to Pfizer. It might be thought that, despite publication in the most prestigious journals and under the apparent authorship of the most distinguished academics, clinicians and researchers would find this literature too obviously industry linked and would not be influenced by it. However, the subsequent citation rates for the Pfizer-linked articles were three times higher than that for the articles on Zoloft not linked to Pfizer (Healy and Cattell 2003).

The profile of this so-called scientific activity suggests that Pfizer ended up with a set of authors whose background increased the possibility of the company's publications appearing in the most prestigious journals. The combination of distinguished journal, distinguished author, efficient distribution system, and sponsored platforms appears to have led to an impact on the therapeutics domain greatly in excess of the impact of the rest of the literature on Zoloft.

At present roughly three-quarters of all randomized trials appearing in *Journal of the American Medical Association*, *New England Journal of Medicine*, or the *Lancet* are industry funded. The impact of this literature on third-party payers is at present unquantifiable, but authorship by perceived opinion leaders with minimal company representation and nondeclaration

of other authorship inputs increase the likelihood that these articles will be influential.

Academics become opinion leaders in a therapeutic field because they have their names on a larger proportion of the literature appearing in the most prestigious journals than their colleagues do, and because they get asked to international meetings to present these data—with which they may not, in fact, have firsthand acquaintance. This, allied with the volume of industry-linked authorship, is arguably leading to a situation in which the dominant figures in therapeutics actually have little firsthand research experience, may have no raw data that they can share with others, and may never have seen the raw data. This is a situation in which, in contrast to the traditional perception of who the ghost authors are in the medical literature, our leading academics have become ghosts or ciphers, valued by companies for their ability to put on an “academic act.”

It is in fact a situation in which ghostwriters increasingly have to take on ghost-acting as part of their repertoire. This happens because the apparent authors of a study will often now have so little familiarity or association with the basic data that they either cannot present it at major meetings or are not inclined to do so in, for instance, poster form. As a result it is becoming increasingly common to find medical writers presenting posters at academic meetings, where they will often be assumed to be doctoral students linked to the research being presented.⁵

The situation that has developed underlines the significance of the proprietary control of raw data. In the CMD series of articles on Zoloft, there were six that dealt with the use of Zoloft for children. Of these six articles, only one mentions suicidality—one single suicidal act. There were in fact six suicidal acts on sertraline in the trials that these articles report: a rate approximately six times higher than the published rate in adults.⁶ The rate of suicidality in depressed children taking sertraline was in fact nine percent. However, the article dealing with the hazards of treatment in children who are depressed only reported on the side effects that occurred at a rate of ten percent or higher (Alderman et al. 1998).

THE CONSENSUS ON TREATING CHILDREN WITH PSYCHOTROPIC DRUGS

The consequences of these developments came to a focus in 2003 on the issue of treating children with psychotropic drugs. The TMAP children’s algorithm project outlined above endorsed the use of SSRI antidepressants for treating childhood nervous disorders, largely on the basis of a series of

unpublished trials. Although unpublished, some of the experts formulating algorithms for the TMAP and later appearing as authors on the few trials published were the same people. These experts therefore had a better opportunity to know what the raw data looked like than anyone else. As a result, the issue of treating children with psychotropic drugs offers a good case example to bring out a number of features of the new world of manufactured consensus.

There has been a long-standing awareness that it is difficult to show in clinical trials that antidepressant drugs offer benefits for children. Despite this there were grounds for using psychotropic drugs for children, and guidelines on the treatment of children who were depressed endorsed such usage (Healy and Nutt 1998). The advent of the SSRI antidepressants offered hope that these agents might be shown to be effective for children where efforts with older agents had failed.

In the early 1990s, regulatory authorities approved the use of the SSRIs Paxil and Zoloft for treating depression in adults, and they subsequently approved Celexa and Efexor. From the 1990s, standard letters of approval to companies noted that as these drugs were likely to be used to treat children, studies to establish the safety of the drugs in these populations would be helpful. This encouragement led to a series of studies of SSRIs in children from the mid-1990s. A further incentive was put in place in 1998 with the FDA Modernization Act (FDAMA) (Sharav 2003), which offered patent extension on the basis of testing for rather than proving safety; if the drugs showed hazards, the company still got patent extension but had to incorporate this information in the label.

Prozac

In the case of fluoxetine an early series of clinical trials failed to establish efficacy in treating childhood nervous problems. A second study that started in 1990 involved extensive prescreening of patients so that fewer than one-fifth of those screened entered the study, and those who did were put through a placebo washout phase in an effort to reduce the high rate of placebo responsiveness found in SSRI trials in children. Using these procedures, an article that appeared in 1997 claimed that Prozac could produce beneficial effects in children and adolescents (Emslie et al. 1997). However, in fact on the primary end-point measure, Prozac was no better than placebo and on secondary measures benefits were apparent on physician-based ratings but not on patient or caregiver ratings. In addition, there was a 29 percent dropout rate on Prozac and the rate of behavioral side effects was higher on Prozac than on placebo.⁷

Subsequently another study had a comparable result (Emslie et al. 2002). The second study, in contrast to both the previous Prozac study and studies of other SSRIs showed no higher rate of adverse events on Prozac than on placebo. This second study had, however, adopted a very unusual procedure, namely, a Prozac test phase, during which any subjects who responded adversely to Prozac were dropped and their data were not included in the final analysis. This combination of studies led to a license for Prozac for the treatment of depression in children and adolescents in 2003.

A further study had been undertaken on Prozac for obsessive-compulsive disorder (OCD). This showed somewhat more clearly positive results for Prozac over placebo but, equally, an excess of suicidality over placebo.

Paxil

The first study undertaken with Paxil, protocol 329, was conducted in the early to mid-1990s. The published report from 2001 pointed to mixed benefits of Paxil on the primary end points of the trial, with apparent responsiveness on some measures accompanied by nonresponsiveness on others, and concluded that Paxil is effective, safe, and generally well tolerated (Keller et al. 2001). But in this study there was an increased rate of suicidal acts on Paxil (5/93; a 5.4 percent rate) compared with either imipramine (1/95) or placebo (0/89). The difference between Paxil and comparators was significant.

These figures were not apparent from the published paper, where suicidal children were coded as having had emotional lability. Hostility was also a reported side effect in 6.5 percent of Paxil patients in this study, versus 1.1 percent on placebo. While the published paper does outline that emotional lability might include suicidal acts, this is not a common meaning of the term for most clinicians, who will be unaware that one dictionary for coding side effects offers the possibility to code suicide, suicidal acts, and suicidal ideation under the heading of emotional lability. The same dictionary codes homicidal acts, homicidal ideation, and other aggressive acts under the heading of hostility.

The second trial, protocol 377, the third trial, protocol 701, and the fourth trial, protocol 716, failed to demonstrate efficacy for Paxil for depression and, also, seem to have returned an increased frequency of suicidality on Paxil. The first two of these studies, which appear to have been completed by 2000, were presented in part in abstracts in 2001 and 2002 that concluded that Paxil was effective, safe, and generally well tolerated (Wagner et al. 2002). The fourth apparently remained unscrutinized by the FDA when the FDA undertook a review of SSRI agents in children in 2003.

At much the same time, studies of Paxil in OCD were instituted, protocols 453 and 704. Reports of these studies in abstract form also claimed that Paxil was effective, safe, and generally well tolerated (Geller et al. 2002; Carpenter et al. 2001). However, company data on file point to an increased rate of side effects on Paxil compared to placebo in the domains of hostility, agitation, and hyperkinesia. In 453, 6.3 percent of children taking Paxil ($n=97$) became hostile, compared with zero percent on placebo ($n=100$). In 704, 9.2 percent of children became hostile on Paxil ($n=98$), with 1 percent becoming hostile on placebo ($n=105$). There was also an increased frequency of suicidal acts on Paxil (1/195) compared to placebo (0/205).⁸

Finally, a study of Paxil was conducted in social phobia, protocol 658. The unpublished results indicate that Paxil might in some cases produce a beneficial effect in children, but as with depression and OCD there was a higher rate of adverse events in the behavioral domain on Paxil compared to placebo. In this case there appear to have been 3 suicidal acts in 165 children on Paxil, compared to 0 in 157 on placebo.⁸

Zoloft

In the case of Zoloft, in the mid-1990s, a double-blind placebo-controlled study was undertaken in OCD, which reported that Zoloft can have a greater beneficial effect on core features of OCD than placebo (March et al. 1998). This paper, which was one in the CMD series, noted one suicidal act on Zoloft. A background expert report on the study, however, points to two suicidal acts on Zoloft, compared with one that was attributed to placebo but did not happen in the randomized phase of treatment.⁹ In the absence of the raw data, outsiders have to take attributions such as this regarding placebo suicidal acts on trust. In the case of Pfizer's clinical trial program in adults over 50 percent of the suicidal acts attributed by Pfizer to placebo occurred during the washout phase of trials and should not have been coded under the heading of placebo (Healy 2003b).

At the same time, Pfizer initiated open trials of Zoloft in children who were depressed. The nine percent rate of suicidality in the Alderman et al. trial, one of the CMD articles, was noted above. A further open study of Zoloft in depression, also in the CMD series, reported that there were three suicidal acts among 53 children who were depressed, a rate of 5.6 percent (Ambrosini et al. 1999).

Subsequently, Pfizer conducted two randomized controlled trials on Zoloft in depression. These were both negative. They were reported in combined form, however, as showing that Zoloft was effective and well

tolerated (Wagner et al. 2003). In fact, 59 percent of children on Zoloft showed a change of five points on a Clinical Global Impression scale, against 49 percent of children on placebo showing comparable changes, a finding that only reaches statistical significance when both studies are combined. In the case of the side-effect profile, there was a doubling of the rate of behavioral problems, including suicidal acts, suicidal ideation, and aggression, in children on Zoloft (6/189) compared to children taking placebo (2/187), and a nine percent dropout rate on Zoloft versus three percent on placebo for adverse events, but in fact 46 of 189 children on Zoloft, 24 percent, dropped out for one reason or another (Garland 2004).

The actual dropout rates on Zoloft contrast with a lower rate of reported behavioral problems in this study compared to earlier studies on both Zoloft and Paxil. In addition, it can be noted that the design of this study did not encourage detection of adverse events. In SSRI studies where side effects are more actively sought, the rates are higher. For example, in a study of fluvoxamine in anxiety, increased motor activity was found in 27 percent of children compared to 12 percent of placebo patients ($p = .06$) (Walkup et al. 2001). This study, in contrast to the Zoloft studies above, used side-effect checklists.

Efexor

In the case of Efexor, two studies have been undertaken in depression and two in generalized anxiety disorder. One study published in 1997 suggests that venlafaxine was safe and well tolerated but that efficacy had not been established (Mandoki et al. 1997). However, it now seems that in the combined depression studies there was an increased rate of children becoming hostile (two percent, versus less than one percent on placebo) and suicidal (two versus zero percent) on venlafaxine compared to placebo (Kuslak 2003). There seems to be no prospect that the full findings from these studies will be published.

THE UNRAVELING OF THE CONSENSUS

In addition to a small number of publications (six full articles with three abstracts) from approximately 15 randomized trials in children, there were approximately 70 publications of open studies or case reports with Celexa, Prozac, Paxil, Zoloft, Luvox, and Efexor. The open studies and published double-blind trials universally portrayed these drugs as being safe, well tolerated, and effective when given to children.

In 2002, the issue of *Newsweek* coinciding with World Mental Health Day carried a cover feature of a depressed teenage girl (Newsweek 2002). The inside story outlined that there were three million depressed teenagers in the United States, and that if left untreated this would lead to a high toll in substance abuse, failed marriages and careers, and deaths from suicide. The article noted that a number of new antidepressants, such as Paxil, Zoloft, and Prozac, could help. Such articles commonly have input from public relations firms working for pharmaceutical companies. The expectation in this case appears to have been that a number of SSRIs would shortly be licensed to treat teenage depression.

It is important to understand what licensing means in this context. It does not mean that physicians are thereafter enabled to treat depressed children in a way that they had been unable to do before. It means rather that Pfizer, Lilly, and Glaxo SmithKline would be enabled to convert the vicissitudes of teenage angst into an illness, one supposedly stemming from a chemical imbalance and one that it was appropriate, indeed almost necessary, to detect and treat.

There are no grounds to believe that the NICE would have come to any different conclusions than the TMAP on the issue of how to treat depressed children, if in due course they had gotten around to considering this issue. Fate and the media intervened to ensure that this never happened.

As a result of a Glaxo SmithKline application to the regulators for a license for Paxil to treat childhood nervous disorders, the raw data from clinical trials were lodged with a number of national regulators. Within a fortnight of seeing the raw data in response to queries as to the events behind the term *emotional lability*, in May 2003 the regulators in the United Kingdom issued a warning against the use of Paxil (Seroxat) for minors. A few weeks later, Glaxo SmithKline wrote to all doctors noting that Paxil use was linked to suicidality and that withdrawal from Paxil was also linked to an apparent doubling of the rate of suicidality. Three months later, Wyeth recommended against the use of Efexor in children, in similar terms. Later that year, in December, the British regulators issued a position statement in which they stated that none of these drugs, bar Prozac, had demonstrated efficacy in depression.

These developments led to a projected FDA hearing for February 2, 2004. Ten days before this hearing, a working group for the American College of Neuropsychopharmacology reported that, after reviewing the evidence, it was the task force's view that SSRI drugs were safe and effective in, and well tolerated by, children (Emslie et al. 2004).¹⁰ The authors of this report included Emslie, Wagner, and Ryan, who had all been authors of study 329 and, among them, had been authors of most of the randomized trial

literature on SSRIs given to children. These three authors and their coauthors, however, noted that they might not be correct in their conclusions that there were no problems with SSRIs in that they had not seen the raw data.

Despite this preemptive strike, in February 2004, an FDA hearing on the use of psychotropic drugs for children recommended strengthening the warnings on these drugs, against a background of regulatory assessments that at least 13 of the 15 studies undertaken of antidepressants in children failed to show efficacy for the drug¹¹ and panel views that there appeared to be an activation syndrome on these drugs.

It transpired that in 1998, a SmithKline Beecham assessment of the early Paxil studies that had been completed at that time, 329 and 377, indicated that the drug did not work for depressed children but that the data would not be submitted to the regulators, as a statement to the effect that the drug had not been shown to work for children would have a negative commercial impact.¹² Selected positive data, however, would be progressed to publication.

What lessons can be drawn from this situation, which probably offers the greatest divide in all of medicine between the raw data on an issue, on the one hand, and the published medical accounts purporting to represent those data, on the other?

First, the data on one side of this divide give the lie to a body of close to 100 papers and abstracts universally reporting the benefits of these drugs lying on the opposite side of the divide. These open and randomized trials, it would seem, have the appearances but not the substance of science. The discrepancy between the papers and the underlying data may stem from the fact that a large number of the key studies appear to have been ghostwritten.

It follows from this that it is almost impossible to accept that these are scientific papers. What the field appears to need is a new term with which to designate such infomercials and a set of criteria that might reliably identify this new genre of marketing product that is the raw material out of which clinical consensus will later be manufactured. This, it should be noted, is the aim of all good marketing—to own the market, not just to sell the product (Applbaum 2004).

A second point is that while pharmaceutical companies know exactly how many prescriptions have been issued and just what each physician writes, almost no one knows how many children or adults are on any psychotropic drugs. When this fact is allied with the fact that physicians report serious adverse events to regulators in no more than 1 in 100 cases, a picture emerges in which Americans and others track the fate of parcels put in the

post 100 times more accurately than they track the occurrence of adverse events on these drugs. The quality of the information reported by patients on adverse events indeed appears to be much better than that reported by physicians (Herxheimer and Mintzes 2004). This is a situation that could not have been tailored better for companies wishing to maximize the expert consensus.

There are grounds to suggest that there must be some fundamental opposition between marketing and science in that the former explicitly operates to build consensus, while the latter supposedly moves forward by fracturing consensus. When we have arrived at a situation in which the mental sets of clinicians have been captured so that it is difficult for them to conceive of alternatives to those being sold to them, there are reasonable grounds to state that such a field is no longer scientific. When there is almost no possibility of discrepant data emerging to trigger a thought that might be unwelcome to the marketing department of a pharmaceutical company, these marketing capabilities would seem to be appropriately described as totalitarian.

IS CHANGE POSSIBLE?

Evidence-based medicine (EBM) is portrayed by its advocates as a value-free approach to the problems of clinical practice. In its early days, the appeal of EBM lay in the promise that the assessment of all available clinical trial data rather than judgments based on selected data sets would deliver clinical facts that should trump the values of individual clinicians, academic or nonacademic, which were all too often at risk of subversion by the free meals on offer from pharmaceutical companies. But as this case study illustrates, there are grounds to think that pharmaceutical companies have effectively subverted the process.

One answer to the difficulties outlined above might be to appeal to the norms of science. Far from being value-free, science values data. “Activities” that involve restricted access to data have not ordinarily been regarded as scientific. We have only recently become aware that our current situation involves a possible breach of the norms of science. This is not a problem that can be tackled by individual clinicians or experts. Ideally the response from clinical societies and journals and independent evaluating agencies, such as the NICE, to this new situation would be to bar pharmaceutical company data from major academic forums or journals or downgrade the value put on these data in developing guidelines, unless companies conform to the

norms of science and make the entire data set from trials available for wider scrutiny.

NOTES

1. In connection with the TMAP, this article has benefited hugely from work undertaken by Allen Jones, Special Investigator in the United States OIG Office of Special Investigations, detailed in *Dwight McKee and Allen Jones v. Henry Hart, Sydni Guido, Wesley Rish, Albert Masland, James Sheehan and Daniel P. Sattelle*, Civil Action No. 4:CV-02-1910, in the United States District Court for the Middle District of Pennsylvania.

2. As of 2004, these guidelines had been adopted at some point by Pennsylvania, California, Colorado, Nevada, Illinois, Kentucky, New Mexico, New York, Ohio, South Carolina, Maryland, Missouri, and Washington, DC, or by jurisdictions in those states.

3. It is important to note that the author participated as a guideline panel member in this Risperdal exercise.

4. It is important to note that the author also participated as a Delphi panel member in this Risperdal exercise.

5. This claim is based on personal experience and discussions with ghostwriters/ghost-actors.

6. Pfizer. 1997. Sertraline Hydrochloride for Obsessive-Compulsive Disorder in Pediatric Patients. Expert Report. New York: Pfizer Inc. Available at: <http://www.healyprozac.com>.

7. FDA review.

8. Data on file. Important Safety Information Regarding Paxil in Pediatric Patients. Glaxo SmithKline, Therapeutic Products Directorate: TDP-Web, July 18, 2003. Health Canada; http://www.hc-sc.gc.ca/hpfb-dgpsa/tpd-dpt/paxil_pa_e.html.

9. Pfizer Expert Report. Sertraline Hydrochloride for Obsessive Compulsive Disorder in Paediatric Patients. Approved October 20, 1997.

10. This was initially available only through GYMR, a Washington-based public relations company that specializes in translating the language of science and medicine into the more understandable language of health. According to GYMR.com, GYMR was

...founded in 1998 by a team of experts in healthcare and social change.... [It] offers clients marketing and communications expertise that strategically support public policy goals.... [Clients] include many of the nation's most respected associations, government agencies, pharmaceutical companies, philanthropic organizations and health initiatives.

Whether it's provoking action on a national health issue or crafting an organizational image that appeals to internal and external audiences, GYMR excels at designing and implementing issue and image campaigns. Our media events are successful because we have a nose for news. We know how to take the language of science and medicine and transform it into the more understandable language of health. We advise clients of the best dissemination strategy for their news and make sure that the message they deliver is compelling, documented and contributes to other national dialogues in a real and meaningful way.

11. <http://www.fda.gov/ohrms/dockets/ac/04/transcripts/4006T1.htm>.

12. Central Medical Affairs Team. Seroxat/Paxil. Adolescent Depression. Position Piece on the Phase 111 Studies. October 1998. SmithKline Beecham Confidential Document, available from the author. This is also available at the *Canadian Medical Association Journal* Web site.

REFERENCES

- Alderman, J., R. Wolkow, M. Chung, and H.F. Johnston
 1998 Sertraline Treatment of Children and Adolescents with OCD or Depression: Pharmacokinetics, Tolerability, and Efficacy. *Journal of the American Academy of Child and Adolescent Psychiatry* 37: 386–394.
- Ambrosini, P.J., K.D. Wagner, J. Biederman, I. Glick, C. Tan, J. Elia, J.R. Hebler, H. Rabinovitch, J. Lock, and D. Geller
 1999 Multicenter Open Label Sertraline Study in Adolescent Outpatients with Major Depression. *Journal of the American Academy of Child and Adolescent Psychiatry* 38: 566–572.
- Appelbaum, Kalman
 2004 *The Marketing Era*. New York: Routledge.
- Carpenter, D.J., G.J. Emslie, B. Birmaher et al.
 2001 Safety of Paroxetine in the Treatment of Children and Adolescents with OCD. Presented at the 40th annual NCDEU meeting, abstract 58.
- Chiles, J.A., A.L. Miller, and M.L. Crismon et al.
 1999 The Texas Medication Algorithm Project: Development and Implementation of the Schizophrenia Algorithm. *Psychiatric Services* 50: 69–74.
- Eccleston, D.
 1993 The Economic Evaluation of Antidepressant Drug Therapy. *British Journal of Psychiatry* 163: Supplement 20.
- Emslie, G.J., J.H. Heligenstein, K.D. Wagner, S.L. Hoog, D.E. Ernest, E. Brown, M. Nilsson, and J.G. Jacobson
 2002 Fluoxetine for Acute Treatment in Children and Adolescents: A Placebo-Controlled Randomized Clinical Trial. *Journal of the American Academy of Child and Adolescent Psychiatry* 41: 1205–1215.
- Emslie, G., J.J. Mann, W. Beardslee, J. Fawcett, A. Leon, H. Meltzer, F. Goodwin, D. Shaffer, K. Wagner, and N. Ryan
 2004 ACNP. Preliminary Report of the Task Force on SSRIs and Suicidal Behavior in Youth. January 21.
- Emslie, G.J., A.J. Rush, A.W. Weinberg, R.A. Kowatch, C.W. Hughes, and T. Carmody
 1997 A Double-Blind, Randomized Placebo Controlled Trial of Fluoxetine in Depressed Children and Adolescents. *Archives of General Psychiatry* 54: 1031–1037.
- Flanagin, A., L.A. Carey, P.B. Fontanarosa, S.G. Phillips, B.P. Pace, G.D. Lundberg, and D. Rennie
 1998 Prevalence of Articles with Honorary Authors and Ghost Authors in Peer-Reviewed Medical Journals. *Journal of the American Medical Association* 280: 222–224.
- Garland, E. Jane
 2004 Facing the Evidence: Antidepressant Treatment in Children and Adolescents. *Canadian Medical Association Journal* 170: 489–491.

- Geller, D.A., K.D. Wagner, G.J. Emslie, T. Murphy, D. Gallagher, C. Gardiner, and D.J. Carpenter
2002 Efficacy and Safety of Paroxetine in Pediatric OCD: Results of a Double-Blind Placebo Controlled Trial. Presented at the 42nd Annual NCDEU Meeting, Session III–16. (Also presented at the APA annual meeting, Philadelphia, May 2002, NR 349.)
- Gilbert, D.A., K.Z. Altshuler, W.V. Rego, S.P. Shon, M.L. Crismon, M.G. Toprac, and A.J. Rush
1998 Texas Medication Algorithm Project: Definitions, Rationale, and Methods to Develop Medication Algorithms. *Journal of Clinical Psychiatry* 59: 345–351.
- Guest, J.F., W.M. Hart, R.F. Cookson, and E. Lindstrom
1996 Pharmacoeconomic Evaluation of Long-Term Treatment with Risperidone for Patients with Chronic Schizophrenia. *British Journal of Medical Economics* 10: 59–67.
- Healy, David
1998 *The Antidepressant Era*. Cambridge, MA: Harvard University Press.
2003a *Let Them Eat Prozac*. Toronto: Lorimer (Linked Web site: <http://www.healyprozac.com>).
2003b Lines of Evidence on SSRIs and Risk of Suicide. *Psychotherapy and Psychosomatics* 72: 71–79.
2004 Shaping the Intimate. Influences on the Experience of Everyday Nerves. *Social Studies of Science* 34: 219–245.
- Healy, David, and Dinah Cattell
2003 The Interface between Authorship, Industry and Science in the Domain of Therapeutics. *British Journal of Psychiatry* 182: 22–27.
- Healy, David, and David J. Nutt
1998 British Association for Psychopharmacology Consensus on Statement on Childhood and Learning Disabilities Psychopharmacology. *Journal of Psychopharmacology* 11: 291–294.
- Herxheimer, Andrew, and Barbara Mintzes
2004 Antidepressants and Adverse Effects in Young Patients: Uncovering the Evidence. *Canadian Medical Association Journal* 170: 487–489.
- Hughes, C.W.
1999 The Texas Children's Medication Algorithm Project: Report of the Texas Consensus Conference Panel on Medication Treatment of Childhood Major Depressive Disorder. *Journal of the American Academy of Child and Adolescent Psychiatry* 38: 1442–1454.
- Keller, M.D., N.D. Ryan, M. Strober, R.G. Klein, S.P. Kutcher, B. Birmaher, O.R. Hagino, H. Koplewicz, G.A. Carlsson, G.N. Clarke, G.J. Emslie, D. Feinberg, B. Geller, V. Kusumakar, G. Papatheodorou, W.H. Sack, M. Sweeney, K.D. Wagner, E. Weller, N.C. Winters, R. Oakes, and J.P. McCafferty
2001 Efficacy of Paroxetine in the Treatment of Adolescent Major Depression: A Randomized, Controlled Trial. *Journal of the American Academy of Child and Adolescent Psychiatry* 40: 762–772.
- Khan, Arif, Shirin R. Khan, Robyn Leventhal, and Walter A. Brown
2001 Symptom Reduction and Suicide Risk in Patients Treated with Placebo in Antipsychotic Clinical Trials. *American Journal of Psychiatry* 158: 1449–1454.
- Kuslak, V.
2003 Letter to Physicians. Wyeth Pharmaceuticals, August 22. Available at: <http://www.rphlink.com/wyethpharmaceuticals.html>.

- Mandoki, M.W., M.R. Tapla, M.A. Tapla, G.S. Sumner, and J.L. Parker
1997 Venlafaxine in the Treatment of Children and Adolescents with Major Depression. *Psychopharmacology Bulletin* 33: 149–154.
- March, J.S., J. Biederman, R. Wolkow, A. Safferman, J. Mardekian, E.H. Cook, N.R. Cutler, R. Dominguez, J. Ferguson, B. Muller, R. Riesenber, M. Rosenthal, F.R. Sallee, K.D. Wagner, and H. Steiner
1998 Sertraline in Children and Adolescents with OCD: A Multicenter Randomized Controlled Trial. *Journal of the American Medical Association* 280: 1752–1758.
- Mortimer, A., D. Healy, R. Gray, R. Peveler, P. Pratt, T. Sharma, and T. Turner
1998 Consensus Statement on Schizophrenia Standards in Care for Maintenance Therapy and Poorly Responding/Treatment Intolerant Patients. CINP meeting, Glasgow. *International Journal of Neuropsychopharmacology* 1:Abstracts Supplement.
- Mosholder, Andrew
1993 Review and Evaluation of Clinical Data. Application NDA 20–272, 11 May. (Cited in J. Whitaker and R. Whitaker 2002. *Mad in America*. Boston: Perseus: 277).
- National Institute for Clinical Excellence (NICE)
2002 Guidance on the Use of Newer (Atypical) Antipsychotic Drugs for the Treatment of Schizophrenia. *Technology Appraisal Guidance* 43, June. Available at: <http://www.nice.org.uk>.
- Newsweek
2002 Depression. 3 Million Kids Suffer From It. What You Can Do. *Newsweek*, October 7:52–61.
- Petersen, Melody
2004 Making Drugs, Shaping the Rules. Big Pharma Is Eager to Help States Set Medication Guidelines. *New York Times*, February 1:3–1, 3–10.
- Rosenheck, R., D. Perlick, S. Bingham, W.L. Mares, J. Collins, S. Warren, D. Leslie, E. Allan, C. Campbell, S. Caroff, J. Corwin, L. Davis, R. Douyon, L. Dunn, D. Evans, E. Frecska, J. Grabowski, D. Graeber, L. Herz, K. Kong, W. Lawson, F. Mena, J. Sheikh, D. Smelson, and V. Smith-Gamble
2003 Effectiveness and Cost of Olanzapine and Haloperidol in the Treatment of Schizophrenia. A Randomized Controlled Trial. *Journal of the American Medical Association* 290: 2693–2702.
- Sharav, Vera Hassner
2003 The Impact of FDA Modernization Act on the Recruitment of Children for Research. *Ethical Human Sciences and Services: An International Journal of Critical Inquiry* 5: 83–108.
- Sheldon, T.A., and G.D. Smith
1993 Consensus Conferences as Drug Promotion. *Lancet* 341: 100–102.
- Wagner, K.D., P. Ambrosini, M. Rynn, C. Wohlberg, R. Yang, M.S. Greenbaum, A. Childress, C. Donnelly, and D. Deas
2003 Efficacy of Sertraline in the Treatment of Children and Adolescents with Major Depressive Disorder: Two Randomized Controlled Trials. *Journal of the American Medical Association* 290: 1033–1041.
- Wagner, K.D., E. Wetherhold, D.J. Carpenter, S. Krulewicz, and A. Bailey
2002 Safety and Tolerability of Paroxetine in Children and Adolescents: Pooled Results from Four Multi-center Placebo Controlled Trials. Presented at the 42nd Annual NCDEU Meeting, Session II–61.
- Walkup, J.T., M.J. Labellarte, M.A. Riddle, D.S. Pine, L. Greenhill, R. Klein, M. Davies, M. Sweeney, C. Fu, H. Abikoff, S. Hack, B. Klee, J. McCracken,

L. Bergman, J. Piacentini, J. March, S. Compton, J. Robinson, T. O'Hara, S. Baker, B. Vitiello, L. Ritz, and M. Roper

2001 Fluvoxamine for the Treatment of Anxiety Disorders in Children and Adolescents. *New England Journal of Medicine* 344: 1279–1285.

World Health Organization

2003 Technology Appraisal Programme of the National Institute for Clinical Excellence. Geneva: WHO. Available at: <http://www.nice.org.uk>.

DAVID HEALY

North Wales Department of Psychological Medicine

University of Wales College of Medicine

Hergest Unit

Bangor, Wales LL57 2PW, UK

E-mail: healy_hergest@compuserve.com