Ethical Human Psychology and Psychiatry, Volume 9, Number 3, 2007

Guidelines, Tramlines, and Faultlines

David Healy, MD, FRCPsych
North Wales Department of Psychological Medicine

M. Savage, MB
R. Tranter, MB
R. Austin, MB
Q. Ijaz, MB
J. A. Hughes, MB
D. Oberholzer, MB
P. Gutting, MB
A. P. Roberts, MB

It would be irrational to want our health care providers to practice in accordance with anything other than the best possible evidence. Given this, it seems entirely sensible to have expert panels regularly assess the current state of the controlled trial evidence for particular treatment options and codify these assessments in guidelines. If these panels are independent, this is surely a good mechanism to ensure that clinicians' practice will be relatively independent of drug company blandishments. We might be able to depend on such a bolster against commercial pressure—if the clinical trials being conducted were genuine scientific exercises and were being reported transparently and honestly. However, recent developments in even independent guidelines raise the prospect that the pharmaceutical industry has captured the guideline process and is using it as a mechanism to gain and control markets. This article aims at bringing out the key issues by considering recent National Institute of Clinical Excellence (NICE) guidelines for pediatric depression and bipolar disorder, and the difficulties such guidelines can pose in certain settings.

Keywords: guidelines; guideline capture; evidence-based medicine; ghostwriting

Most clinicians have been or will be faced with a situation in which the medical director of their employing organization inquires of them individually or as part of a clinical group whether they are compliant with some guideline issued by the National Institute of Clinical Excellence (NICE) or some other body or are compliant with targets set out by government or management.[1] It is not clear what the consequences might be of a failure to comply. A recent medical defense union article, while making it clear that clinicians do not always need to adhere to guidelines, does suggest that any deviation from guidelines needs to be justifiable (Colbrook, 2005). It is also not clear how open health care executives or other managers might be to an argument that a considerable proportion of the material in mental health guidelines or of the targets set for psychiatry are based on less than clear-cut evidence that pursuing these recommendations will in fact lead to better patient care.

Because these things are not clear, guidelines risk becoming tramlines within which clinical practice gets constrained, even though avowedly in most instances they are not supposed to be prescriptive. This article seeks to outline some of the ambiguities in current guidelines, question assumptions about whether guidelines and targets within mental health necessarily serve patients well, and attempts to differentiate between the kinds of evidence bases that should lead to guideline adherence and the kinds that are likely to support fashionable rather than enduring claims.

BACKGROUND

In 2003, Britain's NICE, a body widely regarded as independent of the pharmaceutical industry, was poised to issue guidelines on the treatment of childhood depression. Fate intervened. A few months later it was clear that the literature on which NICE had depended in this area was flawed. It was flawed both because a large number of trials remained unpublished and also because the trials that had been published overemphasized the benefits of treatment and concealed or minimized the hazards. The divide between what the data for this treatment area are in fact now known to show and what the then-published literature claimed the data demonstrated is the greatest known divide of this sort in medicine (Healy, 2006b).

There is no reason to believe that the processes that gave rise to this divide have been confined to the matter of treating pediatric depression. These processes are likely rather to be endemic to psychiatry and probably to a great deal of medicine. Recent controversies involving major medical journals reveal that treatment hazards are systematically downplayed, treatment benefits are consistently oversold, and in the case of articles on therapeutics many of these are ghost-written by authors who have profound conflicts of interest. These articles provide the material on which NICE bases its guidelines.

At one point, following the difficulties with guidelines for the treatment of pediatric depression, NICE considered giving all pharmaceutical company clinical trials a lower banding in the hierarchy of evidence considered in the course of elaborating guidelines. Such a downgrading may never have been politically feasible and has not in fact happened.

Without a mechanism to take into account the distortions outlined above, the guideline process risks capture, with clinicians finding themselves as a result subject to treatment recommendations with which they may profoundly disagree. In recent years, the phenomenon of regulatory capture has been outlined (Abraham, 2002a, 2002b; Abraham & Davis, 2005; Abraham & Reed, 2002). This happens when experts linked to companies sit on regulatory panels assessing the efficacy and safety of drugs, and when regulators depend on company summaries of what company randomized controlled trials (RCTs) show. As a consequence, agencies, whose brief is to ensure on the behalf of the public that only effective agents are licensed and that the hazards of these agents are appropriately labeled, have in some instances licensed treatments of less certain effectiveness and have been much slower to warn of hazards than the public might have expected. They have been captured.

A body like NICE is perhaps even more vulnerable to capture of this sort than are regulatory agencies, in that NICE has no access to the raw data from the trials it considers and cannot access the trials that companies have withheld from publication. The stakes are high in that in Britain getting NICE to endorse a treatment option is possibly the most

efficient marketing step a company can take, and companies within the mental health domain now regularly advertise their products as endorsed by NICE.

BIPOLAR GUIDELINE

Against this background, consider the recent NICE (2006) guidelines on bipolar disorder. These contain a number of generally sensible non–evidence-based suggestions that should be part of standard clinical practice. Where the guideline touches on pharmacotherapy, however, it comes close to having all the problems likely to have been found in the guidelines NICE might have issued for pediatric depression had fate not intervened.

As regards the generally sensible recommendations, these include monitoring the physical health of patients with bipolar disorder. Quite aside from the fact that this is simply good practice, one of the agents NICE recommends for bipolar disorder is linked to a greater frequency of hyperglycemia, type II diabetes, and hyperlipidemia than other agents not mentioned by NICE.

As regards pharmacotherapy, among other recommendations the guideline makes the following five recommendations. First, it emphasizes the use of recent antipsychotics, and does not mention all other antipsychotics that have since 1952 been the mainstay of the management of mania and bipolar disorder. The reason these older agents are not endorsed may be that they do not come with RCT data to support them. The reason for this lack of data is that classic bipolar disorder leading to hospitalisation is relatively infrequent, and when present is typically so severe that it would be difficult to recruit patients to an appropriate clinical trial. But most early papers on chlorpromazine concerned its utility for manic and confusional states, pointing at the same time to its relative inefficacy for schizophrenia (Delay & Deniker, 1952; Delay, Deniker, & Ropert, 1955). No reason has been offered to think these early observations or 40 years of clinical practice were wrong.

But the fact that these older agents do not have RCT data has let pharmaceutical companies seek an indication for newer, probably no more effective, and potentially more hazardous agents in the management of this illness. This has been done by recruiting patients with conditions of lesser severity and perhaps less certain diagnoses to short-term trials that adopt outcome measures that yield some treatment effect that may stem from little more than sedation rather than convincing efficacy. These trials offer the possibility to gain a license for the treatment of the condition. As a result, all of a sudden it has appeared that the only agents supported by evidence for the treatment of mania or bipolar disorder are newer antipsychotics or anticonvulsants.

There are further complexities. NICE suggests using risperidone for the treatment of acute mania, but the biggest trial on which this recommendation is based was conducted in India (Khanna et al., 2005) and has been the subject of a major BBC program ("Drug Trials: The Dark Side," April 27, 2006) that questioned whether patients gave informed consent to the trial and suggested they were essentially being processed through a clinical trial factory. The correspondence in the *British Journal of Psychiatry* on the ethics and validity of this study (Srinivasan, Pai, Bhan, Tesani, Thomas et al., 2006) may be more extensive than for any other study the journal has ever published.

Given that there is a movement of clinical trials to third-world settings (Petryna, 2006), we perhaps face a future in which the bulk of the evidence that might dictate the practice

of psychiatry in Western settings will come from settings that are very different from those in which the treatment will be given and will have been generated in circumstances for which no American or European clinician can vouch. Until relatively recently the trials on which treatment was based had been conducted by relatively local investigators who had hands-on clinical experience with the new treatments and were able to talk authoritatively about the results. This will change if the key data that influences practice stems from other parts of the world, and consequences for clinical practice are uncertain. The most clear-cut consequence is that different ethnic groups have different responses in terms of both efficacy and side effects, but the lack of availability of clinicians who have participated in key trials may also throw up other problems.

Second, in the case of the prophylactic management of bipolar disorder, where NICE recommends the use of one agent, many in the field openly talk about the same trial data (Tohen, Calabrese, Sachs, Banov, Detke et al., 2006) as offering an indication of the likelihood that this drug produces physical dependence and a withdrawal syndrome (Ghaemi, 2005; Healy, 2006a).

Third, NICE recommends stopping treatment with antidepressants after an acute depressive episode has resolved, stating there is no evidence that continuing antidepressant treatment reduces relapse rates. There is, however, little evidence for NICE's position on this issue, and the idea of not giving antidepressant drugs to patients who are very depressed is very clearly an idea of considerable appeal to the marketing departments of companies pushing "mood stabilizers."

Fourth, NICE recommends using valproate for prophylaxis, even though this agent has not received a license for this purpose anywhere. The reason valproate has not received a license for this purpose is because of a lack of evidence supporting its use for this purpose.

Allied to this is the fact that NICE includes a whole series of recommendations involving treatment combinations for patients with frequent relapses or ongoing functional impairment. These recommendations for combined treatment regimens are in general not supported by convincing RCT data.

Few clinicians would have a problem with the recommendations for valproate or for treatment combinations, but advocating these options while failing to mention older agents, which are supported by decades of clinical experience, appears to be endorsing a set of current fashions rather than treatments that have been proven to advance clinical care over alternate options.

Fifth, in its final section on children and adolescents, the guideline does not mention that hitherto unanimous clinical opinion has held that bipolar disorders do not start in childhood. Instead by considering the possibility of treatment for bipolar disorders in childhood, NICE apparently envisages children being given some of the most toxic drugs in use in medicine without any evidence for benefits in the long term. The power of guideline capture can perhaps best be seen in this notion of offering a guideline for the management of bipolar disorder in children. A company does not need to seek an indication for treatment in children if influential guidelines tacitly endorse such treatment.

These latter points need to be read against a background of what appear to be vigorous efforts in recent years in the United States to convert childhood difficulties into diseases to be managed with pharmacotherapeutic means (Harris, 2005). Some experts seem willing to contemplate the diagnosis of bipolar disorder in utero (Papolos & Papolos, 2000). While most European clinicians, perhaps mistakenly, would probably at

present think we are never likely to get NICE guidelines on the management of bipolar disorder in utero, these NICE guidelines are in fact peppered with U.S. diagnostic terminology and clinical terms that until recently few European clinicians would have used.

Finally, there are important omissions. For example, NICE does not include among its priorities for implementation any need to monitor the mental state of treated patients for signs of suicidality, even though current data for the drugs otherwise recommended in this guidelines have been shown in clinical trials to double the risk of suicidal acts compared to placebo (Healy, 2006a; Storosum et al., 2005).

BACK TO GUIDELINES

Many of the problems outlined above stem from efforts to endorse particular practices on the basis of limited data. The same does not apply to guidelines based on studies that point to the inefficacy of treatment options. One of the best instances of this lies in the series of trials that have uniformly indicated that debriefing is not at present an appropriate treatment for posttraumatic stress disorder, leading to NICE guideline recommendations against this treatment option (Bisson, Jenkins, Alexander, & Bannister, 1997; NICE, 2005; Raphael, Meldrum, & McFarlane, 1995). In general, where recommendations are based on clear evidence that particular treatments should not be taken up, it would be foolhardy to deviate from the guideline other than in exceptional circumstances.

Within psychiatry, however misleadingly certain academic papers may be written, with the possible exception of clozapine for treatment-resistant schizophrenia, no body of studies allows claims for a comparative superiority of one pharmacotherapeutic agent over another.[2] The clearest evidence for this lack of superiority lies in the fact that the regulatory authorities have not permitted any company to make claims for comparative efficacy. The studies on which claims are made are all placebo controlled trials, and the limited superiority of these active agents compared to placebo should make it clear that no treatment options currently come close to the kind of evidential threshold that would mandate their use in preference to other available agents. In the absence of compelling evidence, the erection of guidelines that advocate one set of agents over another, however well-meaning, risks producing perverse outcomes.

Within psychiatry, an additional problem involves recognition that through a combination of apparently novel indications and publication strategies, companies can make particular disease areas fashionable, can engineer the appearance of comparative efficacy, and can enlist academic advocates for particular treatment options. If guidelines are going to command widespread support, they will need, however, to reach beyond the fashionable.

If we ask the question whether in the case of an audit of the NICE guidelines for bipolar disorder it would be appropriate to set a particular standard for adherence, it arguably becomes clear that the bipolar guidelines cannot be audited. It is impossible to see how agreement could be reached based on the evidence as to what an appropriate standard might be against which current clinical practice could be audited. This is in complete contrast, for instance, with recommendations regarding penicillin for general paralysis of the insane (GPI).

Perhaps the answer to a medical director or other management inquiry about adherence to this NICE guideline is that we aim to adhere completely with recommendations regarding practices that should not now be undertaken, but as regards other recommendations we await convincing evidence that they are based on something more than current fashion.

NOTES

1. In the case of the authors, the relevant medical director will not be a psychiatrist; she may be a surgeon, and the issue that NICE guidelines might pose some of the problems outlined in this article may seem close to incomprehensible.
2. The situation is no different for psychotherapy. While there is evidence of superiority over placebo, except perhaps for the management of conduct disorder, there is little evidence for a comparative efficacy of one treatment modality over another.

REFERENCES

Abraham, J. (2002a). Making regulation responsive to commercial interests. *British Medical Journal, 325,* 1164–1169.

Abraham, J. (2002b). The pharmaceutical industry as political player. *Lancet, 360,* 1498–1502.

Abraham, J., & Davis, C. (2005). A comparative analysis of drug safety withdrawals in UK and US (1971–92): Implications for current regulatory thinking and policy. *Social Science & Medicine, 61,* 881–892.

Abraham, J., & Reed, T. (2002). Progress, innovation and regulatory science in drug development: The politics of international standard setting. *Social Studies of Science, 32*(2), 1–33.

Bisson, J. L., Jenkins, P. L., Alexander, J., & Bannister, C. (1997). Randomised controlled trial of psychological debriefing for victims of acute burn trauma. *British Journal of Psychiatry, 171,* 78–81.

Colbrook, P. (2005, April 9). Can you ignore guidelines? *British Medical Journal Careers,* 143–144.

Delay, J., & Deniker, P. (1952). 38 Cas de psychoses traitées par la cure prolongée et continue de 4560 RP. *C.R.Congrès Méd Alién Neurol France, 50,* 497–502.

Delay, J., Deniker, P., & Ropert, R. (1955). Etude de 300 dossiers de maladies psychotiques traits par la chlorpromazine en service fermé depuis 1952. *Encéphale,* 528–535.

Ghaemi, N. S. (2005, May). Uses and abuses of evidence based medicine in psychiatry. In *Evidence based psychiatry. What it is and what it is not.* American Psychiatric Association meeting, Atlanta Symposium 37A.

Harris, J. (2005). The increased diagnosis of juvenile "bipolar disorder," what are we treating? *Psychiatric Services, 56,* 529–531.

Healy, D. (2006a). The latest mania. Selling bipolar disorder. *PloS Medicine.* Retrieved from http://dx.doi.org/10.1371/journal.pmed.0030185

Healy, D. (2006b). Manufacturing consensus. *Culture, Medicine and Psychiatry, 30,* 135–156.

Khanna, S., Vieta, E., Lyons, B., Grossman, F., Eerdekens, M., & Kramer, M. (2005). Risperidone in the treatment of acute mania. Double-blind, placebo-controlled study. *British Journal of Psychiatry, 187,* 229–234.

National Institute for Clinical Excellence (NICE). (2005). *Post-traumatic stress disorder. Clinical guideline 26.* Retrieved October 8, 2007, from http://www.guidance.nice.org.uk/CG26/niceguidance/pdf/English

National Institute for Clinical Excellence (NICE). (2006). *Bipolar disorder. Clinical guideline 38*. Retrieved October 8, 2007, from http://www.guidance.nice.org.uk/CG38/niceguidance/pdf/English

Papolos, D., & Papolos, J. (2000). *The bipolar child*. New York: Random House.

Petryna, A. (2006). Globalizing human subjects' research. In A. Petryna, A. Lakoff, & A. Kleinman (Eds.), *Global pharmaceuticals. Ethics, markets, practices* (pp. 33–60). Durham, NC: Duke University Press.

Raphael, B., Meldrum, L., & McFarlane, A.C. (1995). Does debriefing after psychological trauma work? Time for randomised controlled trials. *British Medical Journal, 310,* 1479–1480.

Srinivasan, S., Pai, S. A., Bhan, A., Tesani, A., Thomas, G., Murtagh, A., et al. (2006). Trial of risperidone in India—concerns. *British Journal of Psychiatry, 188,* 489–492.

Storosum, J. G., Wohlfarth, T., Gispen de Wied, C. C., Linszen, D. H., Gersons, B. P., van Van Zwieten, B., et al. (2005). Suicide-risk in placebo controlled trials of treatment for acute manic episode and prevention of manic-depressive episode. *American Journal of Psychiatry, 162,* 799–802.

Tohen, M., Calabrese, J. R., Sachs, G., Banov, M. D., Detke, H. C., Risser, R., et al. (2006). Randomized, placebo-controlled trial of olanzapine as maintenance therapy in patients with bipolar I disorder responding to acute treatment with olanzapine. *American Journal of Psychiatry, 163,* 247–256.

Acknowledgments. Some of the authors of this article would practice almost entirely in line with the NICE guidelines for bipolar disorder mentioned below, others would not, but all believe that a number of the recommendations do not at present have sufficient foundation to constrain practice in the manner implicitly recommended by guidelines endorsement. All of the authors (excepting D. Oberholzer) have been asked to chair or speak at local meetings outlining the evidence in favor of the agents endorsed by these NICE guidelines for both schizophrenia and bipolar disorder.

Correspondence regarding this article should be directed to David Healy, MD, FRCPsych, North Wales Department of Psychological Medicine, Cardiff University, Wales LL57 2PW. E-mail: Healy_Hergest@compuserve.com