The Meaning Response and the Ethics of Avoiding Placebos

Daniel E. Moerman

Eval Health Prof 2002 25: 399
DOI: 10.1177/0163278702238053

THE MEANING RESPONSE AND THE ETHICS OF AVOIDING PLACEBOS

DANIEL E. MOERMAN
University of Michigan–Dearborn

Because the "placebo effect" seems to result from "deception," it is often disparaged and despised. Rethinking this and realizing that these benefits flow largely from the meaning of medical encounters (and are far better understood as "meaning responses"); realizing that there need be no deception to elicit them and that they are often very desirable, engaging fundamental human biological pathways, puts the ethical dilemma in a new light. It seems unethical to avoid—to evade—coming to a full understanding of how meaning can so profoundly improve human well-being.

AUTHOR'S NOTE: Special thanks to members of the Placebo Study Group of the Harvard University Mind/Brain/Behavior Initiative for stimulating conversations about this subject and to Nick Humphrey, Fabrizio Benedetti, Howard Field, and especially to Anne Harrington, who read the manuscript and made incisive suggestions. And thanks to my wife, Claudine Farrand, who also read and made helpful comments on the article.

EVALUATION & THE HEALTH PROFESSIONS, Vol. 25 No. 4, December 2002 399-409
DOI: 10.1177/0163278702238053

In medical lore, placebos—inert treatments—are fraudulent, deceptive, corrosive of medical authority, and therefore to be avoided at all costs unless absolutely necessary. These ideas have considerable historical depth. In a famous passage, Thomas Jefferson—in a letter to a physician in 1807—wrote, "One of the most successful physicians I have ever known has assured me that he used more bread pills, drops of colored water, and powders of hickory ashes than of all other medicines put together. It was certainly a pious fraud." Jefferson is forgiving, but he does credit this action to be a "fraud," even if it was a "pious" one.

More recently, in a very influential 1974 article published in *Scientific American*, philosopher Sissela Bok phrased the problem somewhat differently. Writing in the wake of the emergence of the double blind trial as the "gold standard" of medical evidence, Bok (1974) argued that deceptive practices, like giving inert drugs—placebos—to patients who thought they were active treatments was deeply unethical, a form of lying that would be corrosive of medical authority:

> Honesty may not be the highest social value; at exceptional times, when survival is at stake, it may have to be set aside. To permit a widespread practice of deception, however, is to set the stage for abuses and growing mistrust. (p. 23)

And, more recently, the latest version of the Principles of Helsinki of the World Medical Association (see http://www.wma.net/e/policy/17-c_e.html), says the following, in paragraph C.29:

> The benefits, risks, burdens, and effectiveness of a new method should be tested against those of the best current prophylactic, diagnostic, and therapeutic methods. This does not exclude the use of placebo, or no treatment, in studies where no proven prophylactic, diagnostic, or therapeutic method exists.

Thus, placebo controls are unacceptable in any case in which there is any "proven" treatment (the term "proven" is not defined in the text).

Why this placebophobia? What is it that people mean by placebo? Why is it anathema?

Note first that it is eminently possible to detect what people call "placebo effects" without any apparent placebos in sight. In a recent

study, Amanzio and colleagues showed that, for hospitalized surgical patients, those given analgesics in the normal fashion—visibly by a clinician—reported more pain relief than patients given hidden infusions of the same analgesics via intravenous line (Amanzio, Pollo, Maggi, & Benedetti, 2001). Pain researcher Don Price, in an editorial accompanying the article, said that Benedetti was "assessing placebo effects without placebo groups" (Price, 2001, p. 201). As a first challenge to the received ethical wisdom, note that the patients who received more benefit, a benefit that came from having the experience of the analgesia—what Price called a "placebo effect"—were those who were *not* deceived; those who *were* deceived—who received their narcotics surreptitiously—got less benefit. The issue is less one of deception than of knowledge; in this study, the patients who knew they were getting treatment got more pain relief than the others.

In another study, both placebos and aspirin worked better when they were labeled with a widely advertised brand name than when the same substances were provided in generic tablets. The brand names potentiated both aspirin and placebo (Branthwaite & Cooper, 1981). The knowledge that this was particularly fine aspirin made it work better.

It should be apparent that in neither of these cases did the effects I am considering have anything to do with the placebos. Placebos are, after all, inert, which means they do not do anything. Yet things apparently happened anyway, things not the result of dissembling or fraud. The same aspirin worked better when it was labeled; the same analgesic worked better when administered visibly. It seems clear enough that what made the difference in these two cases was not the fact of treatment but the experience of treatment; in particular, I would argue that the brand name and the visible clinician made the experiences more meaningful.

Thus, I contend that the "placebo effect" is misnamed. It is simply a fallacy to label the therapeutic consequences of being in a control group—one receiving placebos—as the "placebo effect." In any controlled trial, the outcome is a consequence of an assortment of factors, among them natural history, regression to the mean, bias, and so on. Moreover, there may be effects from drugs, and there may be effects that, I suggest, are the consequence of the meaning of the medication or procedure, the construction of understandings shared between doctor, patient, and community, which I call the "meaning response"

(Moerman, 2002). In this view, drugs (active or inert, injected or oral) as well as surgical procedures, diagnoses, prognoses, informed consent forms, and so on can be seen as, among other things, "meaning delivery systems."

Indeed, it seems to me very difficult—absent significant efforts at deception—to avoid conveying meaning to any living patient, even a vaguely sentient one. If such meaning has effects, if such effects are to be labeled "placebo effects," and if placebo effects are to be avoided, then we are in big ethical trouble. Of course, if meaning responses are an inevitable aspect of medicine, and we delude ourselves into imagining that they do not exist or that they only exist when we prescribe inert medication, then we are in, I would suggest, even worse ethical trouble. It is one thing to avoid inert medical treatments; it is simply impossible to avoid meaning in medicine. Let me give a few examples to try to indicate very briefly the range and nature of these inescapable factors.

Dutch researcher Ton de Craen carried out a meta-analysis of 22 trials of sumatriptan for migraine headache. This drug was originally available only in the form of subcutaneous injection. Later it became available in oral form. De Craen compared the effectiveness of various forms of placebo treatment. Subcutaneous placebo was significantly more effective than oral placebo (32.4% vs. 25.7%; 6.7% difference; 95% confidence interval [CI] 2.4%-11.0%) (de Craen, Tijssen, de Gans, & Kleijnen, 2000). It has long been the case in the West that injections have been seen as "more powerful" than pills: thus, a special benefit is "a shot in the arm," not "a pill in the mouth."[1] A meaningful difference like this can affect the outcome of medical practice, and it cannot easily be avoided.

Similarly, many have noted that surgery is powerful and rich in meaning simply by virtue of its nature as invasive and aggressive (Beecher, 1961; Johnson, 1994). An interesting experiment recently showed substantial improvements over baseline conditions for many patients with obstructive hypertrophic cardiomyopathy who participated in a study of pacemakers. Although patients with active pacemakers improved on most measures, so did those who were randomly chosen to have their pacemakers installed but not activated. Active pacemakers achieved marginally better results than inactive ones, but most differences were not statistically significant (Linde, Gadler, Kappenberger, & Ryden, 1999). Note that this is not to say that active

pacing "did not work" because it was not substantially better than inactive pacing. Both conditions led to substantial improvements. This need not lead us to install inactive pacemakers; active pacing produced somewhat better overall results than inactive. But we must recognize that much of the effect of such implements comes from the fact of their insertion, not necessarily from their actual operation in the body—that is, much of their effectiveness comes from the patients' understanding of what the object is, what it does, and, in all likelihood, the physician's faith in it.

This latter point, that the physician's attitude toward the treatment is very important, is seen very clearly in another study, a meta-analysis of trials of drugs for depression. The investigators compared the outcome in early studies of what became "standard" antidepressant drugs (imipramine and amitriptyline) with the effect of those same drugs in later years when they were "comparator" drugs in trials of newly developed antidepressants (amoxapine, trazodine, and malprotiline). When these once new and exciting drugs were bypassed by newer formulations, effect sizes "were approximately only one half to one quarter the magnitude of effect sizes obtained in the [earlier] analyses" (Greenberg, Bornstein, Greenberg, & Fisher, 1992).

I have shown something similar regarding the drug cimetidine, the first H_2 receptor antagonist shown to be effective for peptic ulcer disease. In studies published between 1976 and 1979, 71% of 1,448 patients treated with cimetidine in trials were healed on endoscopy, usually after 4 weeks. After ranitidine appeared in tests in 1980, the effectiveness of cimetidine declined; only 64% of 697 patients were better at the end of those trials ($\chi^2 = 11.18, p > .001$) (Moerman, 2000). Old drugs become less effective when new ones come along. Antibiotics aside, it seems unlikely that this is due to changes in human physiology, the disease, or to changes in the drugs themselves. It is more clearly due to the attitudes of physicians who (at least in those pre-Internet days) were the ones knowledgeable about drugs and their effectiveness.

Pills, although less powerful than injections, are still powerful. They are also discrete items, hence easily countable. So it is not surprising to see evidence from meta-analyses showing that four placebo tablets a day have more effect than two placebo tablets a day (de Craen et al., 1999); in this case, in 79 trials of various treatments for acute duodenal ulcer, 44.2% of patients who received four placebos per day

had fully healed ulcers on endoscopy after 4 weeks, whereas only 36.2% of patients who received two placebos per day were better (difference of 8%, 95% CI 4.6% to 11.3%). Every child past the age of 3 knows that 4 means more than 2.

Note that these factors are simply unavoidable. One might elect to design a trial that eschews inert medications, but if one is to administer medical treatment, it has to be injected, swallowed, inserted as a suppository, rubbed on as a cream, surgically implanted, or whatever. And the fact of administration makes a difference, as does the mode and number of objects, the status of the treatment in the clinician's opinion, and his or her enthusiasm for it.

A common misconception is that these responses to meaning (or to placebos) are not actually real; they are "all in your head." Of course, saying that something is "in your head" is not the same as saying it is not real. I would argue that "language" is "all in your head." What happens in your throat and mouth—the modulation of the breath—is interesting but not central. And two recent imaging studies have shown where in the head two sorts of meaning response occur: in the striatum of patients with Parkinson's disease (de la Fuente-Fernández et al., 2001) and in the rostral anterior cingulate cortex and brain stem in people in pain (Petrovic, Kalso, Petersson, & Ingvar, 2002) but only after a sentence (i.e., "Here's the dopamine" or "This is a powerful analgesic") is decoded in other nearby areas in the brain: in the superior temporal cortex, the angular gyrus, the supramarginal gyrus, the pars opercularis, and several others (see, e.g., Sakai, Hashimoto, & Homae, 2001). The notion that some portions of this symphony of neurons should be suitable for ordinary scientific study but that others should be avoided because they are somehow unethical seems to me simply unethical.

A CASE STUDY

A recent issue of the *Journal of the American Medical Association* (*JAMA*) provides an interesting case study of how these matters work and how they can be confounding to investigators. A major study carried out at Duke University compared the effectiveness of Hypericum perforatum, St. John's wort, and placebo for major depressive disorder. The primary finding of the study was that St. John's wort was not

more effective than placebo; therefore "this study fails to support the efficacy of *H. perforatum* in moderately severe major depression" (Hypericum Depression Trial Study Group, 2002). There is, however, more to the study than that. It included a third study group, this one randomly assigned to receive sertraline (Zoloft). It was also clear in the study that sertraline was not more effective than placebo. Yet all three groups of patients with moderately severe depression—that is, a baseline score of more than 20 on the Hamilton Depression Scale (HAM-D)—showed substantial improvement over the 8-week course of the study, with HAM-D scores declining, on average, about 8 to 10 points. In one measure of full response—defined as Global Clinical Improvement (GCI) score of 1 or 2[2] and a final HAM-D score of 8 or less—placebo-treated patients did significantly better than either herb- or drug-treated patients; regardless, 27% of all patients achieved this level of improvement. Thus, the authors' conclusions seem a bit odd; if Hypericum was not shown in this study to be an effective treatment for depression (because it was not better than placebo), then it seems hard to avoid the conclusion that sertraline was not shown to be effective either. Unless, of course, one recognizes the much more interesting possibility that all three treatments—Zoloft, St. John's wort, and placebo—were more or less equally effective for depression. If nothing else, this study throws up a strong challenge to the Principles of Helsinki; if Zoloft is a "proven" therapy, and therefore this study had been done without the placebo control group, the results would likely have been very differently interpreted.

Another article in the same issue of *JAMA* encourages us in that more interesting conclusion, that the three are more or less equally effective (Walsh, Seidman, Sysko, & Gould, 2002). In a very demanding meta-analysis of 75 controlled trials for depression, researchers showed that, over the past 20 years, the effectiveness of drug treatment for depression has trended up substantially, so that the proportion of patients responding to tricyclic antidepressants and to selective serotonin reuptake inhibitors had increased from about 40% to about 55%. Over the same period, the proportion of patients responding to placebo increased from about 20% to about 35%. The proportion responding was strongly correlated with the year of publication of the study for both drug and placebo treatment. The authors concluded that "some factor or factors associated with the level of placebo response must therefore have changed significantly during this period.

Unfortunately, we were not able to identify these factors" (Walsh et al., 2002, p. 1844).

Over the past generation, there has been a clear shift in consciousness among doctors, patients, friends, and generally everyone to the effect that depression can be treated with drugs. This was simply not the case (or at least not broadly shared) 20 or 25 years ago. As recently as 1970, for example, Goodman and Gilman's *Pharmacological Basis of Therapeutics*, one of the standard reference sources, was clearly more enthusiastic about electro-convulsive therapy (ECT) than it was about treatment with imipramine or amitriptyline, which were said to never be more effective than ECT (Goodman & Gilman, 1970, pp. 186-192).

Today, however, we all "know" that drugs are effective for depression: We read it in the newspapers and in the scientific journals; we see it on television dramas and in drug company advertisements everywhere both in professional media and on TV commercials. Antidepressant drugs are available in the drugstore and, as shown in the Hypericum study, at the drug counter of your local supermarket. This is all quite new.

There has been a shift in what things mean, a shift experienced across the entire culture. As we increase our certainty that drugs can effectively ameliorate depression, they—and their inert mimics—gain efficacy.[3]

Note that the conclusion here is not to replace Zoloft or St. John's wort with placebo tablets. This simply would not work. The key to these studies is that people—patients, family, and doctors (who are, after all, people with families)—know that Zoloft or Prozac or Hypericum "work" (a compelling American metaphor[4]). The more we know this, the more meaningful these objects become to us—the more they do work. There is no deception here. We know what we know.

CONCLUSION

Today, with patients (or are they consumers?) knowledgeable about many prescription medications from widespread advertising on television and in other popular media or from their research on the Internet, these enthusiasms are likely to be more complex than in the

past. It is plausible to predict that, among two similar drugs for the same condition, the one with the larger advertising budget might be expected to be the more effective. Short of a ban on drug advertising (which would certainly founder on First Amendment grounds in the United States at any rate), these factors simply cannot be avoided. They are there, whether we like it or not.

Personally, I like it. Insofar as one drug is more effective because it is more widely advertised than its competitor, it might be feasible to prescribe it in smaller doses to achieve the same end, saving money and perhaps reducing toxic effects (a highly ethical principle, following quickly from "First, do no harm"). Note that, like many other important factors in healing—such as nationality, gender, and ethnic identity—one cannot "randomize" people to these conditions, to preferring this ad to that. We cannot randomly assort part of our study population to know that four is the same as two or to a group that watches a lot of TV and one that does not. What people know about the world is as much a part of their lives as is the fact that they have blue eyes or are left-handed. It is important to recognize that doctors are no different from anyone else on these matters. They also have blue or brown eyes and watch TV. A primary source of medical information and education for most practicing physicians is drug company representatives who provide advertising material for them personally. If Firm A's advertising is more impressive to clinicians than is Firm B's and if this enhances the effectiveness of some drug because it increases their enthusiasm for it, what are we to do? Is the clinician's discussion with the patient a "lie" if she passes on the enthusiasm conveyed to her by a detail man?

Part of the problem here is that we know very little about what is actually happening. Most of the literature that is relevant was originally done for some other purpose, not to illuminate these problems. (There are exceptions; among the most exciting examples have come from Fabrizio Benedetti's laboratory in Turin, some of which he describes in an article elsewhere in this issue.) In a world of placebophobia, we are not likely to see much more of it. But to eschew understanding a fundamental aspect of human biology, the ability of sick people to respond powerfully and positively to meaningful interactions in the clinic, because it is somehow associated with something that may be considered unethical—placebos—strikes me as being unethical itself.

NOTES

1. A recent article in *Nature* is titled "Proposal for a Vaccine Lab Gets a Shot in the Arm" (Knight, 2001).

2. A GCI of 1 means "very much improved," and a GCI of 2 means "much improved."

3. Note that only very few studies of treatment for depression have included an untreated control group, which is essential to know if there is a real "placebo effect" and that improvement of patients was not due to regression to the mean or ordinary fluctuation of disease (natural history). In a study of depression, this would typically involve a "wait list group"; I am aware of two. In one, a study of a behavioral intervention, the Beck Depression Inventory score of the treatment, control, and wait groups were at the beginning of the study 21, 24, and 23, respectively; after 6 weeks, they were 5, 14, and 21. In this study, there was a substantial "placebo effect" and no evidence of regression to the mean or natural history changes (Fuchs & Rehm, 1977). In another study in rural India, an untreated group did not improve over 4 weeks of study, although a placebo-treated group improved after 2 weeks but then declined to baseline conditions (Nandi et al., 1976). It is not clear in this study that all the investigators were blinded to the treatment condition of the patients, which may account for these results. Given the compelling ambiguity about the effectiveness of contemporary antidepressant medication compared to placebo (i.e., Kirsch, Moore, Scorboria, & Nicholls, 2002), it seems absolutely essential to do some methodologically sophisticated three-arm trials for depression and its treatment.

4. In France, for example, drugs do not "work," they "march." One would say "*ça marche bien*."

REFERENCES

Amanzio, M., Pollo, A., Maggi, G., & Benedetti, F. (2001). Response variability to analgesics: A role for nonspecific activation of endogenous opioids. *Pain*, *90*(3), 205-215.

Beecher, H. K. (1961). Surgery as placebo: A quantitative study of bias. *JAMA*, *176*(13), 1102-1107.

Bok, S. (1974). The ethics of giving placebos. *Scientific American*, *231*(5), 17-23.

Branthwaite, A., & Cooper, P. (1981). Analgesic effects of branding in treatment of headaches. *British Medical Journal*, *282*(6276), 1576-1578.

de Craen, A. J., Moerman, D. E., Heisterkamp, S. H., Tytgat, G. N., Tijssen, J. G., & Kleijnen, J. (1999). Placebo effect in the treatment of duodenal ulcer. *British Journal of Clinical Pharmacology*, *48*(6), 853-860.

de Craen, A. J., Tijssen, J. G., de Gans, J., & Kleijnen, J. (2000). Placebo effect in the acute treatment of migraine: Subcutaneous placebos are better than oral placebos. *Journal of Neurology*, *247*(3), 183-188.

de la Fuente-Fernández, R., Ruth, T. J., Sossi, V., Schulzer, M., Calne, D. B., & Stoessl, A. J. (2001). Expectation and dopamine release: Mechanism of the placebo effect in Parkinson's disease. *Science*, *293*(5532), 1164-1166.

Fuchs, C. Z., & Rehm, L. P. (1977). A self-control behavior therapy program for depression. *Journal of Consulting and Clinical Psychology*, *45*(2), 206-215.

Goodman, L. S., & Gilman, A. (1970). *The pharmacological basis of therapeutics*. New York: Macmillan.

Greenberg, R. P., Bornstein, R. F., Greenberg, M. D., & Fisher, S. (1992). A meta-analysis of antidepressant outcome under "blinder" conditions. *Journal of Consulting and Clinical Psychology*, *60*(5), 664-669.

Hypericum Depression Trial Study Group. (2002). Effect of Hypericum perforatum (St John's wort) in major depressive disorder: A randomized controlled trial. *JAMA*, *287*(14), 1807-1814.

Johnson, A. G. (1994). Surgery as a placebo. *Lancet*, *344*(8930), 1140-1142.

Kirsch, I., Moore, T. J., Scorboria, A., & Nicholls, S. S. (2002). The emperor's new drugs: An analysis of antidepressant medication data submitted to the U.S. Food and Drug Administration. *Prevention and Treatment*, *5*(Article 23). Retrieved from http://www.journals.apa.org/prevention/volume5/pre0050023a.html

Knight, J. (2001). Proposal for vaccine lab gets a shot in the arm. *Nature*, *414*(6861), 239.

Linde, C., Gadler, F., Kappenberger, L., & Ryden, L. (1999). Placebo effect of pacemaker implantation in obstructive hypertrophic cardiomyopathy. *American Journal of Cardiology*, *83*(6), 903-907.

Moerman, D. E. (2000). Cultural variations in the placebo effect: Ulcers, anxiety, and blood pressure. *Medical Anthropology Quarterly*, *14*(1), 1-22.

Moerman, D. E. (2002). Explanatory mechanisms for placebo effects: Cultural influences and the meaning response. In H. A. Guess, A. Kleinman, J. W. Kusek, & L. W. Engel (Eds.), *The science of the placebo: Toward an interdisciplinary research agenda* (pp. 77-107). London: BMJ Books.

Nandi, D. N., Ajmany, S., Ganguli, H., Banerjee, G., Boral, G. C., Ghosh, A., et al. (1976). A clinical evaluation of depressives found in a rural survey in India. *British Journal of Psychiatry*, *128*, 523-527.

Petrovic, P., Kalso, E., Petersson, K. M., & Ingvar, M. (2002). Placebo and opioid analgesia—Imaging a shared neuronal network. *Sciencexpress*. Retrieved from http://www.sciencemag.org/cgi/rapidpdf/1067176v1

Price, D. D. (2001). Assessing placebo effects without placebo groups: An untapped possibility? *Pain*, *90*(3), 201-203.

Sakai, K. L., Hashimoto, R., & Homae, F. (2001). Sentence processing in the cerebral cortex. *Neuroscience Research*, *39*(1), 1-10.

Walsh, B. T., Seidman, S. N., Sysko, R., & Gould, M. (2002). Placebo response in studies of major depression: Variable, substantial, and growing. *JAMA*, *287*(14), 1840-1847.