

Examining a Powerful Healing Effect through a Cultural Lens, and finding Meaning

Daniel E. Moerman, PhD*

In this paper I argue that the "placebo effect" doesn't exist; placebos do, but they are inert so they have no effects (that's what "inert" means). Yet we know that often enough, things do happen after placebo administration. Among various causes for such change, I attribute some effects to the meanings the placebos convey to the participants in the medical event—the doctors, nurses, patients, family, community, etc., of the patient. I call these "meaning responses," and survey here some of the ways they occur (with or without the presence of placebos). Then, I describe some recent studies which dramatically complicate the interpretation of RCTs, and our perhaps overly simplistic understandings of the nature of medical efficacy.

INTRODUCTION

In this paper, I will examine a powerful healing effect thru a cultural lens.

The human healing process is complicated, and involves a number of different dimensions, sometime interacting, sometimes apparently orthogonal. The most important component is probably the action of the immune system, operating independently. Also involved are the natural history of many self limiting illnesses (colds, flu, sprains, simple broken bones); a certain amount of conditioning or learning, as we face an illness for the second or third time; regression to the mean (that is, things sometimes just go back to "normal"); bias of patients or investigators trying to please one another and, perhaps, themselves. Medication can play a role, as can *meaning*, that is the cognitive and emotional response to the rich skeins of relationship of objects of thought which are especially lively in times of crisis—like an illness of a spouse, or child, or one's self.

None of these is particularly controversial save perhaps the last; usually understood (or as I would prefer, misunderstood) as the "placebo effect," this important element in the human healing project occasions periodic scorn with reviews denying either that these forces exist at all, or, if they do, they are trivial and short-lasting. I'm sure most readers are familiar with the mini-industry of papers by Drs. Hróbjartsson and Gotzsche which, if nothing else, have drawn attention to some of the very worst papers ever written within the history of medicine (Hróbjartsson & Gøtzsche 2001).

There is a long history of such articles, global attempts to explain away as a delusion one of the most important and interesting forces in human life. Why such "skepticism" persists at the same time as powerful evidence continues to accumulate for the biological consequences of the fact of medical care, it's meaning (rather than its content), is a striking question about a cultural phenomenon. I believe that the answer to that question has at least two parts: first, while reductionism is utterly essential for a scientific approach to human biology, the fact is that some matters are more easily "reduced" than other; some phenomena, like the construction of meaning, are emergent properties of the whole of human mental, emotional, religious, and historical process, and they engage the interactions of people, communities, history and culture, in ways that are as richly interesting and important as they are complex and challenging. Many of these processes are totally and utterly invisible to us at the same time as we embody them; in a sense, I embody meaning in the same sort of way I embody my liver, of which, unless something vary bad is happening, I am ordinarily totally unaware. Confusing these things with related but vastly simpler communication systems (bird song, chimpanzee signing) we trivialize the most astonishing aspects of our being which are so enormous and powerful that we often simply don't recognize them for what they are. We can find and move within such meaning thru an array of cultural processes: rituals, dance, music, literature. Most of it we don't understand, and needn't: none of us save the most skillful and educated linguist can plot out the ins and outs of ordinary language,

*Department of Anthropology, University of Michigan-Dearborn, Dearborn, Michigan, 48128-2406, USA
e-mail: dmoerman@umich.edu

and even the linguists have their limits. Yet we can all talk (well. . . we can try). For a primer on meaning, see Michael Polanyi's book "Meaning" (Polanyi & Prosch, 1975).

Secondly, physicians often find these matters onerous, as it seems to throw even more responsibility onto them for patient outcomes. "My gosh," they say. "Now I'm responsible for my 'bedside manner.'" Easier to ignore the whole thing. To do so, however, is, in my view, to miss something of extraordinary human and medical importance.

This extraordinary will to disbelieve is an interesting and complicated question in itself, but not one we can consider seriously here.

Hróbjartsson and Gøtzsche's study (Hróbjartsson & Gøtzsche, 2001) to the contrary, many studies in both the laboratory and the clinic have shown that people receiving inert treatments have received significant benefit, but few as elegantly and persuasively as those by Fabrizio Benedetti.

In a classic study of experimentally induced pain (Benedetti & Amanzio, 1997), an open injection of saline—presented as a helpful pain reliever in about 6 or 8 words—is given to the members of one group; the outcome is compared to another group which receives a hidden injection of saline—the same injection, but with no words—in the other. That's the only difference between the two groups. Yet the open saline group shows a persistent decline in pain reports while the hidden infusion group shows a continued rise in pain. Let me qualify this: Does this show us that placebos have effects? No, *because both groups got placebos*. The difference between the two groups was *words, language, meaningful utterances*.*

It is not so easy to get such clear evidence of this in the clinic since—largely for ethical reasons—it is difficult to deny sick or injured patients any treatment at all, placebo or otherwise. There are, however, a few such studies, including one looking at third molar extraction.

Gracely's results in a three arm trial with subjects following removal of third molars shows that pain in a placebo treated group declined while in a comparable but untreated group, pain continued to increase for several hours after surgery. Both Visual Analog Scale and Verbal Descriptor Scale pain reports dropped substantially after inert treatment compared to no treatment (Gracely et al., 1979).

One of the biggest difficulties in all this follows from the confusion of what is happening here with the placebo treatment. Imagine that patients in some mythical trial are given inert tablets called placebos. A week later they are different than at baseline; this difference is the "placebo effect." Of course, it's not. Placebos are inert; **they**

don't do anything. One reason people may be different is regression to the mean. Regression to the mean is not caused by placebos, but by study selection criteria (select 1000 people with hypertension; let them alone for 3 months, and many of them will now have "normal" blood pressure as things set themselves right). Placebos don't cause changes due to natural history, and they don't cause conditioning (for conditioning to occur, you have to train the subject with an active drug, one which has an unconditioned response; see, for example (Ader, 1997)). If placebos don't do anything, then it seems possible that what we call "placebo effects" might occur without placebos.

In an important study, 835 women who reported that they regularly treated headaches with over the counter analgesics were randomly placed in 4 groups: one group received unlabeled placebo, one received placebo marked with a widely advertised brand name, "one of the most popular . . . analgesics in the United Kingdom widely available for many years and supported by extensive advertising", one received unbranded aspirin, and one received branded aspirin. They noted the amount of headache pain relief an hour after taking the pills (Branthwaite & Cooper, 1981). Results: First, aspirin was more effective than placebo. But brand name aspirin was more effective than generic aspirin, and brand name placebo was more effective than generic placebo.

In particular, 55% of headaches reported by branded placebo users improved after an hour (rated 2, 3 or 4 on the scale) while only 45% of 410 headaches were reported to be that much better by unbranded placebo users ($P^2 = 6.76$, $p < .01$). Aspirin relieves headaches. But so does the knowledge that the pills you are taking are good ones, which you learned on tv. The difference here is to be attributed not to the placebo (which is, after all, inert) but to the brand name which clearly is not, enhancing the effect of both placebo and aspirin.

Similarly, Benedetti reported on an experiment where surgery patients were treated with four different drugs appropriate to their conditions; however, half the patients received their drugs openly, with an injection by a clinician, while half received equivalent doses of the same drugs by hidden infusion through an intravenous line (Benedetti et al., 2003).

Patients receiving the medication openly, who were told they were about to receive it, reported more pain relief than those who received equivalent amounts of drugs secretly. Pain researcher Don Price, in an accompanying editorial, described this study as "assessing placebo effects without placebo groups." (Price, 2001). As much as I respect Don Price, this is an unfortunate use

* Note that this experiment was a replication on a whole new level of complexity of the pioneering study done a generation earlier by Levine, Gordon and Fields (Levine, et al. 1978)

of language. There were no placebos here. So obviously, there weren't any "placebo effects." What differentiated the separate groups in this study were *human interaction and words*.

Price did, however, recognize this: he noted that although the increase in pain relief in the study was probably not, by itself, clinically significant, "both pain research scientists and the pharmaceutical industry go to the ends of the earth to make improvements of this magnitude [to existing drugs]. Adding one or two sentences to each pain treatment might help to produce them" . Placebos are inert, but language is not!

And a recent study shows that the language need not be deceptive in any way. Kaptchuk has shown that patients with irritable bowel syndrome (IBS) respond favorably to compassionate care, that is, thoughtful and responsive professionals expressing concern, giving accurate information, and carefully discussing concerns with patients. But another group of patients who received the same compassionate care, plus placebo tablets three times a day, which they were told were placebos, that is, inert tablets which had been shown to help people like them in the past, did significantly (and clinically) better than the first group. (Kaptchuk, 2010)

Note that most of the examples I have given deal with pain, clearly the system most fully mapped for meaningful responses. But there are other systems which can also respond to language and meaning. Benedetti has replicated his open/hidden drug experiment in three other areas: diazepam in anxiety state, stimulation of the subthalamic nucleus in Parkinson's patients, and administration of beta-blocker (propranolol) and muscarinic antagonists (atropine) in healthy volunteers . In all these cases, when the treatment was given openly, it was more effective than when given secretly. (Benedetti et al., 2003; Colloca et al., 2004)

THE MEANING RESPONSE

Given that there are no placebos in most of these experiments, it seems unwise to call these responses "placebo effects." And the aspirin study, which shows that the brand name can enhance the effect of an inert drug *and* of an active drug, indicates that at least one dimension of what is going on here is the effect of what medications *mean*. I define the meaning response as "the psychological or physiological effects of meaning in the treatment of illness." Much of what is called the placebo effect—the really *interesting part*, that is, meaning responses elicited with inert medications, is a special case of the meaning response, as is much of what is called the "nocebo effect."

I am interested particularly in the responses that people have to what things mean or to what they know, to what others often call their expectations or expectancies. I don't use these terms since they seem to me as an anthropologist insensitive to culture; I anticipate before the fact that people in different parts of the world with different cultural backgrounds will know the world differently, and might construct different meanings of apparently similar objects or experiences. I would suggest that, more often than not, expectancies are the outcome of a complex play of meanings. The two approaches are not fundamentally different, but have different emphases.

It is also important to note that these matters, where meaning has an influence on health and even mortality, can occur well outside the ordinary bounds of the clinic.

Dr. P.D. Phillips and colleagues have shown that, in the presence of a broad range of diseases in Chinese Americans in California, those who are understood by Chinese traditions of astrology to be particularly susceptible to these conditions—by virtue of the year of their birth—die significantly earlier than those with the same conditions born in other years. Here are two examples from 6 or 8 which Phillips described:

EARTH YEARS

Chinese born in "earth years," that is, years ending with 8 or 9—and consequently deemed by Chinese medical theory to be especially susceptible to diseases involving lumps, nodules, or tumors—*and* who have lymphatic cancer, die, on average, 4 years sooner than Chinese with lymphatic cancer born in other years.

LUNG DISEASES

Those with lung diseases born in "metal years," years ending in 0 or 1—in Chinese theory, "the lung is the organ of metal"—die on average 5 years younger (roughly 7% of length of life!) than those born in other years. There were no such differences found in a similar examination of the mortality of thousands of non-Chinese Californians (Phillips et al., 1993). These are very compelling examples of "meaning responses."

DYING ON THE 4TH DAY OF THE MONTH.

In another study, Phillips showed that Chinese-Americans and Japanese-Americans were more likely to die on the 4th day of the month than any other because 4 is an unlucky number; if 13 is an unlucky number for Californians in general, it's not unlucky enough to increase the mortality rate (Phillips et al., 2001). It is worth noting that these meanings—of metal and the lung, or of earth and

lumps, or of unlucky fours, are not notions concocted by individual patients or therapists; they are icons of a sort which permeate the language and culture of, in this case, immigrant Chinese and/or their American born children, to some degree or other; Phillips shows in one case that the effects of these beliefs are influenced by the degree of commitment to Asian culture. These relationships have nothing to do with having an Asian body, but with having Asian ways of living, thinking, behaving and being.

At least some of the time, biological processes can be “activated,” or perhaps “suppressed,” by that system of meanings we call culture.

Although these effects occur widely in human life, they are often most clearly and visibly displayed in the clinic. People bring to their engagements with physicians many things; patients are not blank slates. But one of the most powerful influences on patients is their doctors. Dozens of studies have demonstrated this; I’ll summarize one.

Such physician attitudes can be conveyed to patients in extremely subtle and delicate ways. Rick Gracely has described a phased experiment in which dental patients were told they would receive either placebo (which might reduce the pain of third-molar extraction, or might do nothing), naloxone (which might increase their pain, or do nothing), the synthetic narcotic analgesic fentanyl (which might reduce their pain, or do nothing), or no treatment at all. Subjects were all recruited from the same patient stream, with consistent selection criteria by the same staff. In the first phase of the study, clinicians (but not patients) were told fentanyl was not yet a possibility because of administrative problems with the study protocol; it is worth noting that fentanyl is

well known in medical circles as a very powerful drug, 100 times more potent than morphine. In the second phase, clinicians were told that now patients might indeed receive fentanyl. Placebo treated patients during the first phase of the study received no relief from it, and, after an hour, their pain reports increased significantly. In the second phase of the study placebo treated patients experienced significant pain reduction from their inert treatment. The only apparent difference between the two groups was that the clinicians knew that no one in the first would get fentanyl while the patients in the second group might (although no one reported on here actually did; they all received only placebo). It is not at all clear how physicians elicited these effects from their patients in a double blind trial. But they did (Gracely et al., 1985); the clinicians were clearly more impressed by fentanyl than were the patients.

The significance of clinician belief, enthusiasm, or commitment, seems to be a fairly broadly applicable principal which can be seen in a number of different contexts.

Old treatments become less effective as new ones come along

It is, for example, a commonplace in medicine that one should use drugs quickly before they lose their effectiveness; this quip has been attributed to William Osler, among others. These data come from a meta-analysis of treatment of ulcer disease (Moerman, 2000). Figure 1 shows the healing rates of drug groups in endoscopically controlled trials of two anti-secretory drugs, plotted by year of publication of the study. At least in the pre-internet world of the 1970s and 80s, it was doctors, not patients, who knew what the hot new drug was. And, apparently, old drugs become less effective as new ones come along.

See figure 1.

Meaning responses occur throughout medicine, in surgery as well as in internal medicine.

I don’t have time to recall here the curious history of the bilateral internal mammary artery ligation which gained popularity in the later 1950s. Suffice it to say that, in two rare double blind trials of a surgical procedure, combined here, people seriously ill with angina and coronary artery disease did as well (maybe a little better!) with the sham procedure as with the “real” one. 75% to 85% of patients experienced substantial subjective and objective improvement increasing exercise tolerance and dramatically reducing nitroglycerine consumption; sham surgery patients received local anesthesia, and two small incisions on the chest which were then closed. These figures are well within the range of improvement for the best contemporary treatments a generation later.

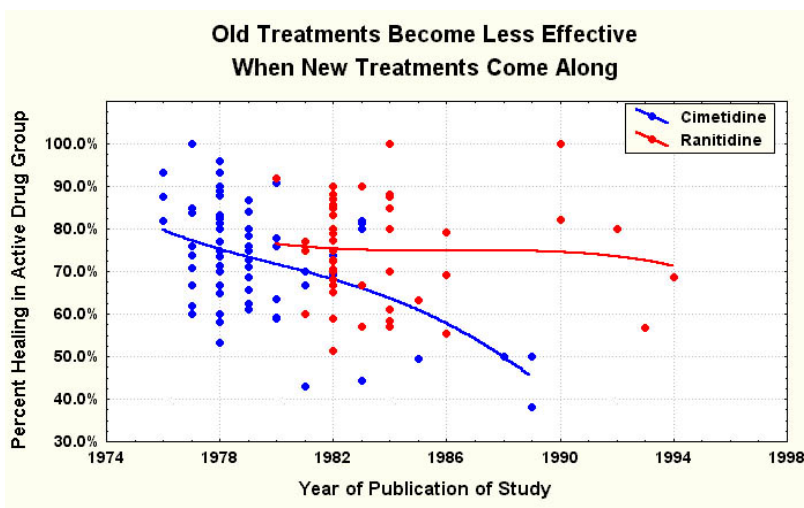


FIGURE 1. Effectiveness studies by date of publication for Cimetidine and Ranitidine, introduced to research in 1977 and 1980, respectively.

mbr.synergiesprairies.ca

A recent study showed the effects of inactive vs. active pacemakers in obstructive hypertrophic cardiomyopathy. Three months after installation of pacemakers, randomly activated or not, all patients were better than at baseline. Sham and active pacemaker patients were better on most dimensions of the study: palpitations, dizziness, shortness of breath, chest pain, self perceived health, and so on. While pacemakers worked better when they were turned on, they weren't much better; they seem to have lowered "Cognitive functioning" which was much improved in inactive pacemaker patients. (Linde et al., 1999)

And in a recent wrinkle in heart surgery, laser transmyocardial revascularization [TMR], there are significant meaning responses as well. Biosense Direct Myocardial Revascularization (DMR) is a variation of this operation in which a laser catheter is inserted in the femoral artery, guided into the left ventricle, and shoots holes in the heart from the inside out.. These surgical procedures are reserved for patients with the most severe and intractable angina. In a quite remarkable study, 299 patients with very serious angina were randomly assigned to high dose, low dose, or no dose of DMR.

At baseline, all patients were rated as class IV on the Canadian Cardiology Society Angina Class (CCS) scale, a physician assessment which was the primary outcome of the study (CSS 4 means "angina at rest, i.e., severe limitation"). Two thirds of the patients improved two or more grades on the CCS. Improvement was substantial, and the same, after 3 months and 6 months for patients who received high dose, low dose, or no treatment with laser catheter inserted but not fired. Disease perception was dramatically improved, as were a broad range of other secondary outcome measures (Leon et al., 2005; Leon, 2000)*.

Alan Johnson made, I think, a prescient observation in 1994; "Electrical machines have great appeal to patients [and doctors] and recently anything with the word 'laser' attached to it has caught the imagination" (Johnson, 1994). He left out the doctors, but I don't! As an aside, I urge you to take a look at the web sites of the companies that make and market these laser instruments, a primary source of patient education for this surgery. Watch very compelling meaning being created out of virtual laser beams with Macromedia Flash.

What we know, what we think, what we are led to believe, or what we understand, whether we know it or not, can have a significant effect in the context of medical care.

Let me add an important caveat; while I believe that it is always prudent to imagine that every medical intervention includes some portion

of the meaning response, it isn't always the case. A study of a statin (Rosuvastatin) and its effect on cholesterol showed a very classic dose response rate: 8 groups of patients were given increasing doses of the drug, from 0 to 80 mg per day. The group with 0 mg had no response. This is unusual, but very interesting; it suggests that the liver operates in a way somehow insulated from neurological influence which seems odd, but not impossible. What this also shows is that we are dealing with a complex form of physiology here, not magic. If it isn't magic, what is it?

I would argue that it is here where we confront one of the biggest and most interesting challenges in this whole arena; and I would suggest that they are very big, very important, and very interesting challenges, Nobel Prize challenges. One of the classic ways that people have dismissed these matters in the past has been by saying, "Well, it's all in your head." It turns out they are right, but not in the dismissive way they intended

Parkinson's disease has long been known by clinicians to be susceptible to influence by inert treatments. Imaging studies by a group from British Columbia have shown a neurological basis for this common clinical observation. Using PET scanning, the authors showed substantial increase in occupancy of D2 receptors with dopamine in the striatum after an injection of saline solution to a Parkinson's patient presented as his standard medication; the increased dopamine crowds out the radioactive dye. These effects are similar in magnitude to the effect of amphetamine in healthy people; the authors note that an area of the nucleus accumbens is also susceptible to placebo effect in Parkinson's (de La Fuente-Fernandez et al., 2001).

In a somewhat more complex study, regional glucose metabolism in PET scans of fluoxetine (Prozac in the US) has been shown to overlap the metabolic pattern of placebo in depressed patients. The active regions in fluoxetine responders overlap the area where activity was evident in placebo responders (Leuchter et al., 2002). Although the clinical response of drug and placebo patients was very similar in this study, drug response in brain activity was somewhat more general than placebo response. In another, similar study, the authors concluded that "Active fluoxetine treatment was associated with additional and unique changes in the brainstem, striatum and hippocampus" (Mayberg et al., 2002). This may help to account for why it is that, while placebo treatment of depression is often very nearly as effective as is treatment with SSRIs, there is often substantially less evidence of unwanted side effects with placebo.

*It may be worth noting that this research, done in the nineties, was first reported at the American College of Cardiology meetings in Florida in 2000. The paper itself was not published for nearly 6 years; reading it shows at least this retired-editor a clear case of a paper dramatically damaged by peer-review.

VARIABILITY IN MEANING RESPONSE

Given this, that it's "all in your head," and that it involves language, it seems reasonable to imagine that cultural factors—different ways of knowing the world through language and meaning—will shape different responses to the same "placebos" around the world. There is a great deal of variability in the response to meaning in medicine, which I wish to look at briefly now.

If there is a single shibboleth in the world of the effect of meaning or of placebos, it is that "placebo effects occur about a third of the time." I can't address the history of this idea here, but I can assure you that it is wrong.

My study of the 4-week endoscopically verified healing rates in 117 control groups in trials of anti-secretory medications prescribed for peptic ulcer disease showed that they ranged from 0.0% to 100% (Moerman, 2000). Meaning responses can be extremely variable. I am convinced that the study of this variability can be a key to developing a fundamental understanding of how meaning interacts with human biology.

COLOUR

Colour makes a difference: changing the colour of a pill can change its effects in a variety of ways. In one particular case, a dozen or more studies have shown that red pills tend to act as uppers/stimulants while blue ones tend to act as downers/sedatives. Moreover, although they probably don't realize they are doing it, drug manufacturers tend to follow suit, colouring their drugs to match these cultural expectations. DeCraen has shown that this is more generally true. He and his colleagues did a study of 49 medicines available for sale in Holland which affect the nervous system. They found that stimulant medications tend to be marketed in red, orange or yellow tablets, while depressants or tranquilizers tend to be marketed in blue, green or purple ones (de Craen et al., 1996)

There are some interesting exceptions to this pattern. In a series of experiments in Italy, it was shown that blue sleeping tablets, or blue placebos presented as sleeping tablets, worked better than did tablets of other colours, but only for Italian women; blue tablets tended to have a stimulating effect on men! Checking with an Italian-American anthropologist colleague*, we came up with this speculation.

Many Italian women have a special relationship with the Virgin who is, in Roman Catholic tradition, the protector of women; in religious art, the Virgin Mary is almost always shown in blue. This iconography—the blue virgin—extends well beyond Italy. But the relationship to

women seems to be particularly strong in Italy although it may be so elsewhere as well. What about men. "Azzuri" is the name (and the colour) of the Italian national football team. Blue, for many Italian men, is not a colour of solace but of excitement and stimulation, of joy and madness, of exhilaration and, too often, of catastrophe (The Italian team last won the World Cup in 2006, the last quadrennial games, beating France.) But it's hardly the colour of sleep.

Symbols are often polysemic; a single colour, the same blue, can be associated with stimulation and excitement, and also associated with solace and protection. It seems a plausible way to think about the experimental results. Anecdotally it may be worth mentioning that Italian football fans holler "Forza Azzurri" to cheer on their team. Let me note that I have a certain sympathy for this proposition since a good translation of "Forza Azzuri" is "Go Blue,"[†] the chant of the University of Michigan Wolverines who wear maize and BLUE colours as they routinely play (another sort) of football in Michigan Stadium in front of 110,000 screaming fans. Not a soporific sight, even for a jaded fan like me. It is also true that the French national football team, the World Cup Champions in 1998, are known as Les Bleus; I am aware of no evidence to show that this has any effect on sleeping tablets in that country, or in Michigan, for that matter.

FORM

As colour can make a difference in the meaning of medicine, so can form. For example, Ton de Craen has shown that injected placebo is more effective than oral placebo in the treatment of migraine headache (de Craen et al., 2000). When the drug sumatriptan (known in the United States as Imitrex, and as Imgran elsewhere) was first introduced, it was only available in the form of an injection; today it is still available that way, but also as tablets and nasal spray. De Craen did a meta-analysis of 35 trials. In placebo treated patients, among those treated with a pill taken by mouth, after two hours 26% of patients reported that their headache was better (it was gone, or mild). Of those treated with a placebo injection, 32% of patients were better. This difference is small (6.7%) but it is statistically significant ($P^2 = 9.4, p = 0.002$).

NUMBER

Similarly the number of pills can make a difference. In a very subtle meta-analysis, Ton DeCraen showed that in some 80 studies of several antisecretory medications for peptic ulcer, there was a significant difference in the endoscopically verified healing rates for those who took two

* Lola Romanucci-Ross suggested this idea.

† Friend and colleague—and native speaker of Italian—Emanuela Appetiti contributed to this interpretation, and assisted with the translation of the highly idiomatic "Forza Azzurri."

placebos per day (36%) compared to those who took 4 per day (44%), a difference of 8 percent, where $P^2 = 21.7, p < 0.0000$ (de Craen et al., 1999).

NATIONAL CULTURAL DIFFERENCES

There are also other cultural factors which are associated with some variation (in addition to colour). Recall the study of inert injection vs. inert tablet for migraine; shots worked better than pills. In studies which were carried out in the United States, the same pattern appeared: 22% oral vs. 34% subcutaneous placebo relief rate. In studies done in Europe however, the difference disappeared: 27% oral vs. 25% subcutaneous placebo relief rate (de Craen et al., 2000; Moerman, 2002) Injections work better than pills, but only in the USA. There are cultural differences shaping the placebo effect.

In my work with peptic ulcers, the mean placebo healing rate (again, 4-week, endoscopically controlled) in 6 German studies is 59%; the rate in 3 Brazilian studies is 7%. For ulcer patients taking placebo, the NNTb (the number needed to treat for a benefit*) for being German (not Brazilian) is 2. There is no obvious reason why this should be the case; to contextualize these differences would require a challenging and complex research study (for which I have not been able to get funding after years of effort).

Sticking closer to home for the Germans: Perhaps Brazilians and Germans have fundamentally different kinds of ulcers (although there is no evidence for this at all). Comparing the 6 German studies to 5 studies from Germany's northern low-country neighbors in Denmark and the Netherlands, the German placebo healing rate is 59% compared to the Danish/Dutch rate of 22%. For ulcer patients taking placebo, the NNTb for being German (not Danish or Dutch) is 3.

Note that the situation is not a simple one; these differences seem to vary by illness: the control group healing rates in treating *hypertension* are substantially lower in Germany than in other Western nations (Moerman, 2000). These are not generic cultural phenomena, or "racial" phenomena, but seem to be specific cultural ones as different conceptualizations, or understandings, or constructions of illness in different cultures seem to have a real impact on health and healing.

HISTORICAL VARIATION

Indeed, these kinds of differences can be seen not only between different cultures, but through time as attitudes and understandings change. Walsh reviewed 75 trials of various antidepressants: tricyclics, and SSRIs compared with placebo. The

effectiveness of drug treatment for depression has trended up substantially, so that the proportion of patients responding to tricyclic antidepressants and to SSRIs 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 conclude 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).

However, the matter doesn't seem too complicated to me. 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 never to be more effective than ECT (Goodman & Gilman, 1970).

Today, while we practically never hear of ECT[†], we all "know" that drugs are effective for depression; we read it in the newspapers, in the scientific journals; we see it on tv dramas, and, in the US at least, we see it in drug company advertisements everywhere, both in professional media and on tv commercials, blogs, and, of course, in our spam e-mail plus Twitter and Facebook.

Antidepressant drugs are available in the drugstore, and, in the form of St. John's Wort, at the drug section of your local supermarket[‡]. As we change our views of the effectiveness of drugs, their effectiveness changes, as do their placebo mimics in trials. Meanings change and so do meaning responses.

And although I have yet to find any indication of it in the medical literature, there are press reports (Silberman, 2009) which suggest that recently many large drug companies, after losing millions of dollars in drug development after their investigational drugs failed to outperform placebo in Phase II trials have become very concerned. There are indications that they have formed a secret committee to compare all their own trial results to determine what is happening, why the effect of dummy drugs is increasing.

* The NNTb, or number needed to treat is a clinically relevant statistic describing the value of a drug or treatment. If in a trial with 100 in each of two groups, and all of the drug group got better while none of the control group did, then the NNTb, the number needed to treat to get one person better is 1. If (more realistically) 50 of the drug group and 25 of the control group improved, the NNTb would be the inverse of the relative risk ($1/(0.5 - 0.25) = 1/0.25 = 4$); that is, you would need to treat 4 people to get one person better.

† It is also true that, whatever public opinion might have to say about ECT after "One Flew over the Cuckoo's Nest," it remains a valuable treatment of choice for many with the most severe and debilitating depression (Kelly and Zisselman 2000)

‡ In a complex RCT, *Hypericum* was shown to be more effective than sertraline (Zolof); however both were less effective than the placebo control. All three treatments led to quite satisfactory outcomes for about a third of patients treated (HDTSG 2002).

CONCLUSIONS

What we know, understand, think, and feel; what we are told and believe; our cultural background; the relationships we have with our clinicians—our doctors, residents, interns, nurses, aids, orderlies, and probably receptionists and parking lot attendants—can very directly affect our response to

medicines, inert or otherwise. These matters are, these days, largely left to chance, or to ideology, or to market forces, but are rarely subject to robust science, although that's less true today, thankfully, than it was a decade ago. The clinical implications of these matters are clearly rich and full, and virtually unexamined.

AN "APPENDIX": PLACEBO DILEMMAS, OR, MEANING STRIKES BACK.

In the fall of 2009, the NEJM published two remarkable articles in the same issue describing two randomized controlled trials, one from the Mayo Clinic, and one from a group of investigators in Melbourne, Australia. The trials looked at a surgical procedure called vertebroplasty of painful osteoporotic vertebral fractures. Older persons, especially women, often have osteoporosis which leads to a weakening of the vertebrae which sometimes simply break. With vertebroplasty, the broken vertebra are repaired with an injection of medicinal glue—polymethylmethacrylate or PMMA. There are an estimated 750,000 persons in the US with such fractures; there are as many as 9 vertebroplasty procedures per 1000 persons in the US annually, and the annual direct care for these fractures in the US is estimated to range from \$12 to \$18 billion in 2002 (Weinstein 2009). This is, then, a substantial industry.

In the Australian study, 71 patients finished the trial; the verum group had the standard procedure while the control group underwent a sham procedure where no needle was inserted into the bone. In the Mayo Clinic trial, 131 patients completed the trial; both groups received anesthesia in the fractured vertebra, but the control group did not receive the subsequent injection of PMMA. In both studies, PMMA was opened and released in the operating chamber since it has a distinctive odor (although it would probably not be all that distinctive to patients).

In the Mayo Clinic study, at one month after surgery, "there was no significant difference between the vertebroplasty group and the control group. . . [and] both groups had immediate improvement in disability and pain scores after the intervention" (Kallmes, et al. 2009). In the Australian study, "there were significant reductions in overall pain in both study groups at each follow-up assessment. At 3 months, the mean reductions" in overall pain score

were the same for both groups. "Similar improvements were seen in both groups with respect to pain at night and at rest, physical functioning, quality of life and perceived improvement" (Buchbinder, et al. 2009). In both studies, all the patients got better, with active or sham surgery. It is important to note the dilemma here: ordinarily, if the treatment isn't better than the control, it is abandoned as a treatment. But here, both the treatment AND the control made significant improvements in the lives of elderly patients (treatment: 74.2 yrs old; control: 78.9 yrs old), mostly women.

Complicating matters are two recent studies of acupuncture for low back pain, one done in Germany (the GERAC trial), and one done at several sites on the US west coast. In the GERAC trial, 1162 patients participated in a three-arm trial: traditional Chinese acupuncture in one group, sham acupuncture (superficial needling at non-acupuncture points in the second group, and conventional care (drugs, physical therapy, and exercise following the standard German guidelines) in the third. Outcome was measured in terms of response to one of two standard pain questionnaires. The results at 6 months showed that 48% of the verum acupuncture group was better; 44% of the sham acupuncture group was better, and 27% of the standard care groups was better. "Low back pain improved after acupuncture treatment for at least 6 months. Effectiveness of acupuncture, either verum or sham, was almost twice that of conventional therapy" (Haake, et al. 2007).

The American trial enrolled 638 adults who were randomly assigned to 4 groups: individualized acupuncture, standardized acupuncture, simulated acupuncture (using a toothpick in a guide tube), and standard care (at the patient's and their physician's choice, usually a combination of drugs and physical therapy). Primary outcome was based on telephone interviews

based on the Roland–Morris disability Questionnaire; there were a range of secondary outcome measures. The patients in the acupuncture groups all gained 4.4 to 4.5 points on the disability scale compared with 2.1 points for the usual care group 8 weeks after treatment. These changes generally were still evident, in the same pattern, after 52 weeks. "In conclusion, acupuncturelike treatments significantly improved function in persons with chronic low back pain" (Cherkin, et al. 2009).

These remarkable studies demonstrate just how effective "inert" therapy can be. But of course the bone surgery in each group was complex and dramatic; there's no room here to describe the elaborate fluoroscope suites where such work is done (many of these procedures are done by radiologists!) There are elaborate and technology-rich diagnostic procedures involved that narrow down a complex aching, debilitating pain to a simple, visible, line on an x-ray, repairable with a little drop of super glue.

Likewise, acupuncture has led a charmed life in America since James Reston's Chinese appendectomy in 1971. Millions of Americans have had acupuncture treatments though perhaps only a half dozen of them understand anything about qi. And now we know that acupuncture is better for treating low back pain (a major nemesis of conventional medicine) than is ordinary medical care, but it is not substantially better than sham treatment.

How to reconcile these studies with our everyday understanding of medical causality is challenging. But they do force us to consider the role—throughout all of medicine—of the power of medicine's ritual, narrative, and performance, which, more and more often, seem to trump its "evidence based" remedies.

REFERENCES

- Ader, R. (1997). The role of conditioning in pharmacotherapy. In A. Harrington (Ed.), *The placebo effect: An interdisciplinary exploration* (pp. 138–165). Cambridge, MA: Harvard University Press.
- Benedetti, F., Maggi, G., Lopiano, L., Lanotte, M., Rainero, I., Vighetti, S., & Pollo, A. (2003). Open versus hidden medical treatments: The patient's knowledge about a therapy affects the therapy outcome. *Prevention & Treatment*, 6, Article 1. Retrieved from <http://psycnet.apa.org/journals/pre/6/1/1a.html> doi: 10.1037/1522-3736.6.1.61a
- Benedetti, F., & Amanzio, M. (1997). The neurobiology of placebo analgesia: from endogenous opioids to cholecystinin. *Progress in Neurobiology*, 52(2), 109–25.
- Branthwaite, A., & Cooper, P. (1981). Analgesic effects of branding in treatment of headaches. *British medical journal (Clinical research ed.)*, 282(6276), 1576.
- Buchbinder, R., Osborne, R. H., Ebeling, P. R., Wark, J. D., Mitchell, P., Wriedt, C., . . . Murphy, B. (2009). A randomized trial of vertebroplasty for painful osteoporotic vertebral fractures. *New England Journal of Medicine*, 361(6), 557–568.
- Cherkin, D. C., Sherman, K. J., Avins, A. L., Erro, J. H., Ichikawa, L., Barlow, W. E., . . . Deyo, R. A. (2009). A randomized trial comparing acupuncture, simulated acupuncture, and usual care for chronic low back pain. *Archives of Internal Medicine*, 169(9), 858–866.
- Colloca, L., Lopiano, L., Lanotte, M., & Benedetti, F. (2004). Overt versus covert treatment for pain, anxiety, and Parkinson's disease. *Lancet Neurology*, 3(11), 679–684.
- 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., Roos, P. J., Leonard de Vries, A., & Kleijnen, J. (1996). Effect of colour of drugs: systematic review of perceived effect of drugs and of their effectiveness. *British medical journal (Clinical research ed.)*, 313(7072), 1624–1626.
- 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-Fernandez, 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.
- Goodman, L. S., & Gilman, A. (1970). *The pharmacological basis of therapeutics*. New York, NY: The Macmillian Company.
- Gracely, R. H., Dubner, R., Deeter, W. R., & Wolskee, P. J. (1985). Clinicians' expectations influence placebo analgesia. *Lancet* 1(8419), 43.
- Gracely, R.H., Deeter, W. R., Wolskee, P. J., Wear, B. L., Sayer, J. S., Heft, M. W., . . . Sweet, J. (1979). The effect of naloxone on multidimensional scales of postsurgical pain in nonsedated patients. *Society for Neuroscience Abstracts*, 5, 609.
- Haake, M., Müller, H. H., Schade-Brittinger, C., Basler, H. D., Schäfer, H., Maier, C., . . . Molsberger, A. (2007). German Acupuncture Trials (GERAC) for chronic low back pain: Randomized, multi-center, blinded, parallel-group trial with 3 groups. *Archives of Internal Medicine*, 167(17), 1892–1898.
- Hypericum Depression Trial Study Group. (2002). Effect of Hypericum perforatum (St John's Wort) in major depressive disorder. *The Journal of the American Medical Association*, 287(14), 1807–1814. doi:10.1001/jama.287.14.1807
- Hróbjartsson, A., & Gøtzsche, P. C. (2001). Is the placebo powerless? An analysis of clinical trials comparing placebo with no treatment. *New England Journal of Medicine*, 344(21), 1594–1602.
- Johnson, A. G. (1994). Surgery as a placebo. *Lancet*, 344(8930), 1140–1142.
- Kallmes, D. F., Comstock, B. A., Heagerty, P. J., Turner, J. A., Wilson, D. J., Diamond, T. H., . . . Jarvik, J. G. (2009). A randomized trial of vertebroplasty for osteoporotic spinal fractures. *New England Journal of Medicine*, 361(6), 569–579.
- Kaptchuk T. J., Friedlander E., Kelley J. M., Sanchez M. N., Kokkotou E., Singer J. P., . . . Lembo A. J. (2010). Placebos without deception: a randomized controlled trial in irritable bowel syndrome. *PLoS One*, 5(12), e15591.
- Kelly, K. G., & Zisselman M. (2000). Update on electroconvulsive therapy (ECT) in older adults. *Journal of the American Geriatrics Society*, 48(5), 560–566.
- Leon, M. B., Kornowski, R., Downey, W. E., Weisz, G., Baim, D. S., Bonow, R. O., . . . Kuntz R. E. (2005). A blinded, randomized, placebo-controlled trial of percutaneous laser myocardial revascularization to improve angina symptoms in patients with severe coronary disease. *Journal of the American College of Cardiology*, 46(10), 1812–1819.
- Leon, M. (2000). DMR in regeneration of endomyocardial channels trial (DIRECT). *American College of Cardiology*. Retrieved from <http://www.acc.org/education/online/trials/aha00/direct.htm>
- Leuchter, A. F., Cook, I. A., Witte, E. A., Morgan, M., & Abrams, M. (2002). Changes in brain function of depressed subjects during treatment with placebo. *American Journal of Psychiatry*, 159(1), 122–129.
- Levine, J. D., Gordon, N. C., & Fields, H. L. (1978). The mechanism of placebo analgesia. *Lancet*, 2(8091), 654–657.
- Linde, C., Gadler, F., Kappenberger, L., & Ryden, L. (1999). Placebo effect of pacemaker implantation in obstructive hypertrophic cardiomyopathy. PIC Study Group. Pacing In Cardiomyopathy. *American Journal of Cardiology*, 83(6), 903–907.
- Mayberg, H. S., Silva, J. A., Brannan, S. K., Tekell, J. L., Mahurin, R. K., McGinnis, S., & Jerabek, P. A. (2002). The functional neuroanatomy of the placebo effect. *American Journal of Psychiatry*, 159(5), 728–737.
- 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). *Meaning, Medicine and the "Placebo Effect"*. Cambridge: Cambridge University Press.
- Phillips, D. P., Liu, G. C., Kwok, K., Jarvinen, J. R., Zhang, W., & Abramson, I. S. (2001). The Hound of the Baskervilles effect: Natural experiment on the influence of psychological stress on timing of death. *British medical journal (Clinical research ed.)*, 323(7327), 1443–1446. doi: 10.1136/bmj.323.7327.1443
- Phillips, D. P., Ruth, T. E., & Wagner, L. M. (1993). Psychology and survival. *Lancet*, 342(8880), 1142–1145.
- Polanyi, M., & Prosch, H. (1975). *Meaning*. Chicago, IL: The University of Chicago Press.

- Price, D. D. (2001). Assessing placebo effects without placebo groups: An untapped possibility? *Pain*, *90*(3), 201–203.
- Silberman, S. (2009, August 24). Placebos are getting more effective. Drugmakers are desperate to know why. *Wired*, *17*(9). Retrieved from http://www.wired.com/medtech/drugs/magazine/17-09/ff_placebo_effect?currentPage=1
- Walsh, B. T., Seidman, S. N., Sysko, R., & Gould, M. (2002). Placebo response in studies of major depression: Variable, substantial, and growing. *The Journal of the American Medical Association*, *287*(14), 1840–1847.
- Weinstein, J. N. (2009). Balancing science and informed choice in decisions about vertebroplasty. *New England Journal of Medicine*, *361*(6), 619–621.