Statistical Challenges in Medical Research: What Consumers Need to Know

Ronald R. Gauch, Ph.D.
Marist College

Abstract:
This paper explores the statistical features inherent in clinical trials that make it so difficult to come up with the right answers. No matter how skilled and diligent the research team, their trial may still produce the wrong conclusions. Consumers, healthcare professionals as well as the general public should be aware of the liabilities that threaten clinical investigations. In spite of the fear and trepidation usually associated with statistics, the issues discussed in this paper can be readily understood because they do not involve complex formulae or esoteric terminology. Facing up to the intrinsic handicaps of medical investigations can mitigate overly optimistic beliefs held by consumers and generate a greater appreciation for the critical role statistics plays in the design and execution of clinical trials.

Key Words: Medical research, clinical trials, human error, statistical issues

1. Clinical Trials and Human Error

The clinical trial is the backbone of medical progress. It is responsible for many of the important advances in treating dreaded diseases such as cancer. Survival rates for breast, uterine, prostate and bladder cancer all improved because of clinical trials. Clinical trials showed that breast cancer could be treated just as effectively with limited surgery as with major surgery, sparing patients unnecessary suffering and disfigurement. Clinical trials debunked the myth that a synthetic estrogen was useful to prevent miscarriages when in fact it caused more harm than good in women trying to maintain a pregnancy. Clinical trials showed that vaccines could prevent a wide variety of dreadful diseases – smallpox, diphtheria, tuberculosis, etc. Clinical trials identified effective drugs to treat debilitating illnesses like rheumatoid arthritis and muscular sclerosis. I can go on and on extolling the extraordinary successes of the clinical trial. There is no doubt that clinical trials can provide a valid answer to an important medical question. However, it can also end up with erroneous findings and the reason this is so needs to be articulated. Consumers, healthcare professionals as well as the general public should be aware of the liabilities that threaten clinical investigations.

A clinical trial is vulnerable to error because of mistakes made by researchers and because of inherent liabilities that even the best researchers can not overcome. I'm mainly interested in the inherent problems associated with a clinical trial, primarily because they are all based on statistical principles. However, I'll begin by first touching on the human mistakes that produce incorrect medical findings.

` The researchers may design a trial that uses the wrong treatment doses, outcome measurements or just doesn't last long enough. The statistical analysis of a trial may be inappropriate, the underlying assumptions of the statistical test may not be met or there is a failure to control the risk of a type I error because of the number of analyses performed. There can be ethical breeches such as permitting unacceptable patients into a trial or entering fake data for missed observations and tests. There's the possibility of recording errors and a failure to ask follow-up questions that could influence the validity of a patient's subjective response to a question. Studies could be designed and treatment results purposely interpreted based on a political or religious ideology. Publication pressure of time and the limited size of a manuscript can cause the omission of useful information in medical journal publications. Preliminary results may be touted at scientific meetings before the data are fully analyzed and cautiously interpreted.

The public's perception of medical triumphs and disasters is often distorted because what the public learns about medical research does not come directly from the research community. What they learn about medical treatments comes from the popular media – newspapers and TV shows in particular. Even when medical researchers share their concerns and point out flaws in their research practices with their colleagues, these limitations rarely trickle down to the millions of people who prescribe, dispense or use medications. The public is frequently awed by what is involved - biology, chemistry, pharmacology and statistics – they fear their lack of understanding about these subjects means they could never understand the research process. The media are also handicapped due to the restricted amount of time or space they are given to translate the scientific information that makes it appealing to the public. In the end the public version of a medical research finding can be markedly different than the real truth.

Note that the events described above that lead to erroneous medical results are derived from human errors. The mistakes can be made by a broad cast of professionals that make up a research team. The errors these individuals commit are due to many causes - a lack of information, too few resources, inadequate training, economic pressures as well as carelessness, sloppiness, greed and arrogance.

Is it any wonder that in 2005 a prominent medical researcher, J. Ioannidis, writing in *PLoS Medicine* declared that most clinical research findings are false. His argument for the high percentage of incorrect medical findings included statistical issues such as small studies, small differences between treatments, the excessive number of tests preformed and the misuse of p values. However, he also pointed out that other factors such as difficulties with outcome definitions, loose trial designs as well as conflicts of interest on the part of investigators also contributed to spurious results.

What perhaps was even more startling was the response by the medical community to this audacious claim. There was deafening silence. In the six months following his article, there were three short responses, published by the same journal that published the original essay. Although the commentators each found Ioannidis' contention provocative, they basically offered helpful suggestions regarding the issues he raised while accepting his basic premise. The low level of criticism or support is in contrast to the fact the article apparently was of high interest with over 100,000 downloads from the publisher's Web site. The article even made it into the popular press with a *Boston Globe* editorial referring to it as a "cult classic". Perhaps the lack of criticism was due to the emphasis on statistical issues. The editors of the journal that published the original article noted that parts of the paper were based on assumptions that even they did not fully understand.

There was a little more reaction the following year as three other articles appeared, but only one challenged the conclusions of the confrontational criticism. Two articles essentially accepted the original suppositions and expanded on its ramifications. The contrary article agreed that there were more false claims in medical research than many would believe were present, but they challenged the over 50 percent figure. They felt that the over estimate could be traced to a flawed mathematical model that Ioannidis had used to "prove" his point. I personally do not believe that the dour position of the Ioannidis paper is true and I suspect that there are many who share my more sanguine conclusion. Nonetheless, there is no question that it is impossible to know for sure if a clinical trial has come up with the right answer.

Don't assume that the problems with clinical research must lie with the individuals conducting the studies as tempting as that may be. Some of the most dedicated, smart and hard working scientists perform medical research. They are ethical, inventive and inspired professionals. The trouble is that the process they must use is inherently flawed. The things they must know are unknown. The things they must control are uncontrollable.

2. Randomized Clinical Trials

The most potent clinical trial is the randomized controlled trial commonly referred to as an RCT. I'd prefer to see a B inserted in the acronym to represent blinded. You can have an unblinded randomized and controlled trial, but unless blinding was impossible, few would consider it a satisfactory research approach. However, whereas the RCT is the best we have, it is not perfect. There are seven inherent problems with an RCT that can corrupt any study and there's little investigators can do about them except hope that none of the flaws infect their studies. I like to call these impediments the seven deadly flaws – it's kind of catchy, but keep in mind these are not equivalent to the seven deadly sins that disciplined early Christians. They are not flaws of commission; they are flaws beyond the control of medical researchers.

What is exceptional about the seven flaws is that they are all grounded in statistical principles. And of even greater merit is that they can all be conveyed and understood by people who never had to take a statistical course, hated the required statistical course they were required to take, or took a statistical course, did well, but remember little of what they learned. To appreciate why medical researchers can not guarantee they have the right answer from a clinical trial does not require a knowledge of statistical significance, type I and II error, the alpha and beta risk or a single statistical formula. The flaws are embedded in the heart and soul of the clinical trial and although the conscientious researcher might be able to reduce the risk of a flaw, he or she can't provide complete immunity. The seven flaws are:

1. The unknown population – the process of selecting patients for a clinical trial is inappropriate.
2. The imperfect sample – only volunteers can participate in a clinical trial and they may create an unrepresentative sample.
3. The inequality among treatment groups – randomization can not guarantee equality.
4. The uncontrolled experimental setting – there are many factors beyond the investigator's control that can bias a trial.
5. The breakdown of blinding – it's too easy to find out who is getting what treatment.
6. The impractical result – clinical trial settings are dissimilar to medical practice and their results may not be reproducible in the hands of your personal physician.
7. The insufficient sample size – you can never have enough subjects. Even the largest clinical trial is too small to find the deadly but rare adverse effect.

2.1 The Unknown Population

The theory that grounds clinical trial research is based on the following paradigm. (a) Define the population, (b) Draw a representative sample from the population, (c) Do a research study on the sample, and (d) Infer your results from the sample back to the population

Note, it all begins with a precise definition of the population. The goal of research is to make statements about a population based on study results from a sample. It is important to know that population because it is suppose to be the source from which the sample is to be drawn. The whole idea of inferential research by using a sample to represent the entire population depends upon an accurate identification of the population.

That's the theoretical model. In medical research that means that out there - somewhere - are all the people who a new drug is intended to help. In statistical jargon, that group of people is the "population". For example, a population could be all the bald-headed men in Virginia. Now it's clear we cannot treat all those people – and, of course, statistical theory doesn't require that we do that. But we are suppose to take a sample from the population that we want to help. Now we add a significant clause to the challenge. Statistical practice requires that we take the sample in such a way that all individuals in the population have an equal chance of being selected. Of course, in order to do that we must first identify all the people in the population.

Here's the troubling part: in medical research we don't and can't adequately identify the population of interest. We don't know and have no way of knowing all the bald headed people in Virginia. Take the illnesses for which we want cures – cancer, heart disease, AIDS, Alzheimer's

disease, etc. We cannot identify the entire population for these diseases. We can't even come close. In addition to the sheer size of a population, the definitions of diseases are not all that specific and people with a disease are often undiagnosed.

Contrast this situation with a poll that is taken to see whom might win the forthcoming election in your town. Only registered voters vote and a list of registered voters exists and can be readily obtained. The list of registered voters is the population. It is clearly the group you want to make inferences about. You can, with relative ease, identify people on those lists and produce a sample. Now you might not get them all to participate in your poll, but at least you have a legitimate sample from that population. Medical research is messier. In medical research, it can't work that way because of the unknown population.

But let's be realistic. Theory is one thing and practice is another. You're never going to identify all the people in the U.S. or in Virginia with a given disease. If you could identify many of them, and then assume that the ones you missed were similar to the ones you identified, you'd be in very good shape. However, even that's next to impossible to do, and in medical research, investigators must settle for less.

Remember the research model requires us to take a sample of individuals from the population, but that's not how it's done in medical research. We don't start by picking patients – the process begins by finding researchers and we're now heading down the wrong path. The patients that researchers have access to are a unique group of individuals. They are not representative of all patients that have the disease of interest. A trial that begins with selecting researchers can easily end up with a set of atypical patients when we contrast that set to the set called for – the population. The results of such a trial cannot tell us for sure, how a treatment will work in the general pool of patients who have the disease.

Usually the researchers selected are the ones who are most interested in treating the targeted disease. He or she may apply for a grant to do a study and add other researchers who share an interest in the disease and who have access to patients. Sometimes a governmental agency may decide to sponsor research for a disease and it solicits researchers at important medical facilities to apply to be the experimenters; these individuals may have a convenient but specialized group of patients. Private parties such as pharmaceutical companies may also contact individual physicians or medical centers and ask those with available patients to participate in a medical study. There are many ways to recruit investigators, but note how all the possibilities violate the process of selecting patients from the population of those with the disease.

There is clearly a problem here – the limited population derived from physician recruitment is obviously different from the theoretical population we should be concerned about. However, how serious is this discrepancy? It all depends on how different the limited population used in the clinical trial is from the true population of interest. If the two populations are very similar, the results of a clinical trial could be very relevant and trustworthy. But without making a studied comparison, we do not know if that level of comparability between the populations is present or not. Any trial may have this problem and that means the results from any trial may not apply to the true population of interest.

Amazingly, this qualification is rarely noted and researchers make broad generalizations just as if the legitimate population had been accessed. Researches publish their results for a drug with a great deal of fanfare, but only infrequently are we forewarned that those results may apply only to the type of patient treated at the research centers used in the study rather than the general assortment of patients in the population.

2.2 The Imperfect Sample

The second inherent flaw that afflicts clinical trials also has to do with the patients selected for the study. To become a participant in a clinical trial you first must volunteer for the study. You can't be a participant in a trial until you give your written consent. The problem that

arises is whether volunteers are the same kind of people as those who want no part of a RCT. Those who won't sign up for a clinical trial may be less desperate than those eager to find a treatment that may help them. Those who forsake a clinical trial will tend to be healthier as well. On the other hand, volunteers also are more likely to be risk takers – after all in a RCT, they take the chance that they will receive the control treatment rather than the more promising experimental treatment. The population of interest has both kinds of patients, those who volunteer and those who won't volunteer, and there is evidence that the two groups have different personal and health characteristics. A 2003 article in *Perspectives in Biology and Medicine* reviewed three investigations that had looked into medical conditions of volunteers who participated in early clinical pharmacology trials. The authors concluded that a large proportion of the volunteers had a history of psychiatric illness, plus other medical conditions and temperaments that differed from the general population.

If, the differences between volunteers and non-volunteers have an effect on the outcomes measured in a clinical trial, then there is clearly a problem because the results apply only to the volunteers. Under this scenario, the clinical trial results are not relevant to people who would not choose to participate in a clinical trial. Unfortunately, we are stuck with this potential dilemma because ethical standards demand that only individuals who volunteer and give their consent can be used in a clinical trial. Consequently, it's not possible to known whether the responses by volunteers and non-volunteers are similar or different and we are left with the possibility that the findings from a trial are not relevant to people who want no part in a clinical investigation.

2.3 Unequal Treatment Groups

The best kind of medical study has a number of important attributes. It needs to include a control group and the make-up of the control group needs to be similar to the treatment group. Typically, the control group receives a placebo, the experimental group receives an active medication and the results from the two groups are compared. So far so good, but any group differences are valid only if a critical condition is met. Namely, at the start of a clinical trial the two groups should be equivalent in terms of the critical variables that affect the trial's end points. For example, we certainly wouldn't want to find that almost all of the sickest patients ended up in the same treatment group. Therefore, in a clinical study subjects are assigned to the control and treatment groups so any inequality between the groups is minimized. Today the acceptable method to do this is to use a random process to make the treatment assignment.

Statistical theory gives us some protection when it comes to factors that confound the results, but it is hardly sufficient. First of all, the theory tells us that the larger the sample the more likely there will be group equivalence. So, if the study is large enough, the process of randomization will distribute many confounding factors fairly equally between the treatment groups. But there is no certainty that this will occur for all important factors. Using conventional standards there is only a 5 percent risk that a variable will be disproportionately represented. However the greater the number of baseline factors that could bias the results, the greater the likelihood that randomization will not provide overall treatment group equivalence. Consequently, in spite of randomization researchers will inspect the distribution of critical baseline variables (e.g. severity of illness) to see if there is an unusual distribution among the treatment groups. If there is an important difference an analysis may be done that adjusts the data so the inequality for the offending baseline variable is neutralized. If the adjusted analysis produces a finding that is inconsistent with the unadjusted result, cautionary comments about how the findings should be interpreted are provided.

Unfortunately, this strategy can not be implemented if the researchers are unaware of baseline factors that could have an effect on an outcome variable. For instance, investigators cannot be aware of all the risk factors for a disease if some are not known when the study is conducted. As an example, the Human Genome project is discovering all sorts of connections between our genetic makeup and a propensity to develop a disease or respond to a treatment for a disease. In 2004 the *New York Times* reported that scientists had made a discovery that surprised even them. They discovered a genetic variation that could predispose people to depression. The

presence or absence of this gene could explain why some people respond to a certain antidepressant and others don't.

Obviously, failure to account for the disproportionate presence of this gene among treatment groups could distort results from an antidepressant study. Previous research findings from antidepressant trials could be at risk as well. Perhaps there was an unbalanced distribution of subjects with this gene that caused a positive or negative result, which was mistakenly attributed to one of the drugs employed in the trial.

2.4 Uncontrolled Experimental Setting

The manner in which a trial is conducted is a concern that transcends many disciplines involved in a clinical trial, but it is the statistician who is usually the strongest advocate for protecting a trial from outside factors. The statistician's heightened preoccupation with controlling the environmental factors of a trial comes from a concern about increased variation and not just the possibility of introducing treatment bias.

The classical model of an experiment requires that all conditions, except for the treatments being administered, to be the same for the groups being tested. This condition simply cannot be met in a clinical trial. You cannot restrict human beings so they behave the same way and have identical environmental exposures. There is enormous variation in terms of how men and women choose to live, what they eat, how much they exercise, the amount of stress they endure, etc. Human beings live in a broad array of environments that places unique pressures and demands on them.

Look at a physical science such as physics or chemistry and compare their research environment to that of clinical medicine. A physical science provides all sorts of ways to produce identical experimental conditions that are impossible to replicate in the clinical setting. In the physical sciences all the relevant variables can be held constant (heat, light, temperature). We can also move from basic science to that of a biological laboratory experiment and appreciate the inferior setting of a clinical investigation.

Consider a laboratory experiment testing the effect of a diuretic versus a placebo in rats. Select ten rats and randomly allocate them to the diuretic or placebo group. The rats are inbred, identically reared and handled. Furthermore, they do not vary significantly in their pharmacological responses. They are housed, fed and manipulated in an identical manner. The biologist compares the outcome variable, urine output, between the rats in the two groups. If the diuretic is effective, there will be no overlap in the result for the groups. All the rats receiving the diuretic will excrete more urine than any of the placebo rats.

This answer comes about because all relevant variables are held constant and the results in the diuretic and placebo treatment groups will either be very similar (if the diuretic is no good) or distinct (if the diuretic is effective). When conducting trials, control of all relevant variables is a goal in clinical research, but it is unreasonable to expect a researcher to even come close to that objective. No matter how hard researchers try, holding constant all the relevant environmental variables of a clinical trial is beyond their control. For example, during the life of a study, events are happening in the environment that can influence how a subject in a study responds. Assume newspapers and TV channels report that a class of drugs, which is being investigated in the trial, causes major side effects or has a positive effect on a person's libido. Such news could easily influence the reports given by the participants in the study. Furthermore, there can be a cold or heat spell in the middle of a trial causing unique responses by subjects that would not occur under more temperate weather conditions. Quite simply, idealized experimental conditions are not possible in clinical research. This does not mean that researchers cannot get a truthful answer, but it does mean they have to be lucky to avoid all the pitfalls that are lurking in the shadows. When it comes to conducting clinical trials even the best researcher cannot overcome an uncooperative environment.

2.5 Breakdown of Blinding

The failure to keep hidden which subject is getting which drug can poison judgments as well as other assessments, and represents the fifth threat that can undermine a clinical trial. Good experimental practice requires that subjects be handled in the same fashion so that the effect of inconsistent patient treatment does not jeopardize the results. Knowing what they are taking also must be hidden from subjects so they don't let that knowledge bias their responses.

It must be realized that blinding is not always possible. For example, in a trial comparing surgery with cancer chemotherapy, blinding would be impossible. But more importantly, in a clinical trial the blind can be broken unintentionally. There is the possibility that an active drug will produce a unique reaction (e.g. dry mouth). Subjects who experience that reaction and associate it with the active drug have clearly broken the blind. Subjects may also recognize a medicated state, particularly when they have received an active drug for the illness being studied in the past.

Patient feedback to an investigator can also defeat blinding. Even if the clues given off by a drug are subtle, without realizing it the investigator may become aware of the treatment a subject is on. For example, the effects produced by an active medication may clearly expose its identity when there is a clear sign (e.g. flushing) present in a large proportion of the cases taking an active drug, but absent among the placebo-control subjects. An illustration of the extent of what can go wrong is demonstrated by a study performed at the SUNY Health Science Center in Syracuse, NY. In 20 of the 23 psychotropic studies examined the authors found evidence that both clinicians and patients knew well beyond chance whether real drugs or placebos were being administered. It should be pretty obvious that the degradation of a trial because blinding failed is a real possibility and successful blinding should never be taken for granted.

2.6. The Impractical Result

The goal of most clinical trials is to make a statistical inference by using the result from a sample to tell how well a drug will do in a larger population of patients. However, think of all the elements that researchers use to control the research environment: not allowing the "wrong" concomitant agents, using only the patients that are most likely to respond to treatment, demanding that the subjects take the treatments as directed to name a few examples. We now end up with a paradox. Many of the factors that researchers introduce to make the RCT tight and protect it from unwanted biases turn out to contribute to a major disadvantage – you may not be able to generalize the results beyond the restricted clinical trial environment.

When we get a positive result it's possible that it applies to an almost unrealistic situation because of all the restrictions placed on a clinical trial. How do we know if the garden variety of patients will do as well as the highly selective ones used in a RCT? Note the many differences between real life and the rarified setting of a research trial. Do patients who forget to take their medications fare as well as subjects constantly prodded to take their trial medications faithfully? Do patients who see their doctors once a year do as well as subjects who are seen weekly? Do patients who eat poorly and rarely exercise do as well as subjects who are on a strict diet and exercise program? The typical RCT has an idealized setting and the observed result may not be conferred upon a more laissez faire setting.

An interesting study on this topic appeared in a 1998 report by a research team at Duke University. They looked at the patients currently having bypass surgery using a special technique. Only four percent of them would have satisfied the selection criteria that were used in the trials that justified the new procedure. It's probable that most of the patients receiving the special technique were very similar to those used in the clinical research phase and were unlikely to be at any significant risk. Nonetheless, this example shows that there can be a sizeable discrepancy between clinical research and clinical practice

2.7 An Insufficient Sample Size

The final flaw is yet another blow to our high expectation for medical research accuracy - the concern over the number of subjects in a clinical trial. It stands to reason that the more observations you make, the greater the assurance of an accurate overall assessment. Take too few observations and you may miss finding something important. Concerns about the number of observations (i.e. number of subjects) are especially relevant to clinical trials.

The number of subjects for a clinical trial can be determined by a formula, but that calculation may require information that a researcher does not possess and it applies to only one variable measured in a trial. However, a typical clinical trial involves scores of tests for a broad assortment of variables and for these assessments, the sample size selected for the main variable may be too small or too large.

Having a sufficient number of subjects is especially difficult when doing research that involves major outcomes such as life or death. A vast sample size is also necessary to identify a rare but perilous side effect. In these situations, researchers usually need an enormously large number of patients, five thousands or more, and time to complete such trials, five years or more. If studies are not large enough it is likely that the answer generated may be due to chance. But even large sample sizes may not overcome all the threats that keep researchers from coming up with the right answer from a clinical trial.

Even outcomes of mega-trials (i.e. trials containing 1,000 or more subjects) can give inconsistent results. A review article in the *Journal of Clinical Epidemiology* described 289 pairs of mega-trials, in which each pair contained the same treatment and type of subject. For example, the two trials would be identified that had the same kind of subjects (e.g. patients with elevated cholesterol) and the identical test treatments (e.g. the same active medication and control treatment). The study conclusion about whether the active treatment was better than, equal to or worse than the control treatments were then compared. A judgment was made on whether there was consistency (did both trials find the active treatment better) or inconsistency (did one trial show no treatment difference but the other trial conclude that there was a significant difference between the treatments). In spite of the enormous number of subjects in these trials, the results of 79 out of the 289 pairs, or 27%, produced inconsistent results. What do we conclude? Even when different trials research the same question, and use very large sample sizes there can be inconsistent trial conclusions; some trial have come up with the wrong answer.

3. Conclusion

As I conclude this talk, it is important to point out that although the probability of a single flaw distorting a finding is small, escaping all the flaws is a matter of luck. Nevertheless, individually or collectively the seven flaws represent serious threats to the integrity of any study. They stand as reminders that as good as the clinical trial is, it may not be good enough. Still in spite of all these threats from all these places, good results do surface and each of us owe the many research teams that conduct medical investigations a vote of appreciation for their fortitude, perseverance and perhaps a bit of divine intervention in their search for the right answers.

Bibliography

Brody J. (2002). Ferreting for facts in the realm of clinical trials. New York Times. Oct 15, 2002 p F7.

Fisher S, Greenberg R. (1993). How sound is the double-blind design for evaluating psychotropic drugs? J Nerv Ment Dis. 181;345-350.

Furukawa T, et al. (2000). Discrepancies among megatrials. J Clin Epidemiol. 2000:53;1193-1199.

Gauch R. (2000). Statistical methods for researchers made very simple. University Press of America.

Gauch R. (2008). *It's Great! Oops, No It Isn't. Why Clinical Research Can't Guarantee The Right Medical Answers*. Springer.
Goodman S, Greenland S. (2007). Why most published research findings are false: problems in the analysis. *PLoS Med* :4;e168.
Hlatky M, Califf R, Harrell F, et al. (1988) Comparison of predictions based on observational data with the results of randomized controlled trials of coronary artery bypass surgery. *J Am Coll Cardiol:* 11;237 -245.
Pauker S.(2005). The clinical interpretation of research. *PLoS Med*. 2;e395.
Ioannidis J. (2005). Why most published research findings are false. *PLoS Med*:2;e124.
Ioannidis J. (2005) Author's reply. *PLoS Med* 2005:2;e398.
PLoS Medicine Editors. (2005). Minimizing mistakes and embracing uncertainty. *PLoS Med*. 2;e272.
Shaywitz D. (2006). Science and shams. *Boston Globe*, July 27.[1] .
Shrier I. (2005). Power, reliability, and heterogeneous results *PLoS Med* 2;e386.
Tishler C, Bartholomae S. (2003). Repeat participation among normal healthy research volunteers. Professional guinea pigs in clinical trials? *Perspect Biol Med* 46;508-520.

[1] www.boston.com/news/science/articlesa/2006/07/27/science_and_sham/