Evidence-based medicine: Four fundamental problems with the randomized clinical trial (RCT) used to document chemical medicine

Søren Ventegodt, MD, MMedSci, EU-MSc-CAM, Niels Jørgen Andersen, MSc, Bernard Brom, MB ChB (UCT), CEDH (France), Dip Acup, Joav Merrick, MD, MMedSci, DMSc, and Donald E Greydanus, MD

1 Quality of Life Research Center, 2 Research Clinic for Holistic Medicine and 3 Nordic School of Holistic Medicine, Copenhagen, Denmark, 4 Scandinavian Foundation for Holistic Medicine, Sandvika, Norway, 5 Intern University College, Graz, Austria, 6 Private practice, Stellenbosch, South Africa, 7 National Institute of Child Health and Human Development, 8 Office of the Medical Director, Division for Mental Retardation, Ministry of Social Affairs, Jerusalem, Israel; 9 Kentucky Children’s Hospital, University of Kentucky, Lexington, and 10 Pediatrics and Human Development, Michigan State University College of Human Medicine, Kalamazoo, Michigan, United States of America

Abstract: Randomized clinical trials (RCT) have been accepted as the golden standard of testing, thus making chemical medicine “evidence based”. The RCT is based on four assumptions: 1) The placebo effect is represented by a placebo pill, 2) it is possible to make a double-blind test with biologically active drugs, 3) beneficial and harmful effects of drugs are fairly measured in RCTs, and 4) an appropriate time frame for the test is used. We have found problems with these assumptions: 1) The placebo effect provided by close relationships to a physician is stronger than an inert pill, 2) double-blind tests cannot be made with biologically active drugs, as these leave an internal clue in the patient that destroys the blinding (active placebo), 3) lack of global outcome measures makes toxic effects invisible for the test and magnifies minor effects to make clinically insignificant positive effects look important, and 4) RCTs are used in such a brief time frame that side effects and harm are not properly detected. The four errors combine into a serious error: The RCT-procedure induces a strong bias in favor of any toxic drug tested. RCTs can turn drugs that are only toxic and not beneficial at all into products sold as useful chemical medicine. Many pharmaceutical drugs on the market today are tested only with this flawed RTC-procedure, and we recommend that these drugs be tested again using a rational method. If drugs are not more helpful than placebo, then we should return to classic psychosocial holistic medicine.

Keywords: RCT, placebo, outcome measures, health, quality of life, evidence based medicine, holistic medicine, adverse effects, adverse events

Correspondence: Professor Joav Merrick, MD, MMedSci, DMSc, Medical Director, Division for Mental Retardation, Ministry of Social Affairs, POBox 1260, IL-91012 Jerusalem, Israel. E-mail: jmerrick@zahav.net.il

Received: September 01, 2009. Revised: October 25, 2009. Accepted: November 01, 2009.
INTRODUCTION
For more than 2000 years, ever since 400 BCE, European medicine was identical to the proud tradition of Hippocrates’ psycho-social medicine. Hippocrates medicine has also been called “character medicine” as its interventions served the purpose of helping the patient to self-insight into his or her purpose of life and character, i.e. total pool of physical, mental and spiritual talents (1).

When a person use all talents to create value in all personal relationships, this person is happy, healthy, and well functioning in all major areas of life, from social, family, and sexual life to working life. The classical physician was a man of wisdom and his training was about helping people to develop self-insight and realize their full potential—i.e. to “step into character”. This treatment intervened only on consciousness and the healing effect of it was called placebo, from the Latin “I please”. The ethics of Hippocratic medicine was the famous principle of primum non nocere—first do no harm (1).

Around 1950, the discoveries of penicillin and other effective pharmaceuticals started a biochemical revolution. If medicine could be drugs, this was an eminent business opportunity, and soon a large number of commercial pharmaceutical companies were established. The first problem the new chemical medicine had to solve was technical chemical problems: how were bioactive drugs with specific effects designed and developed? The second problem was how such potentially poisonous and harmful drugs could be tested on humans? The third problem was ethical—to convince the medical societies to abandon the traditional Hippocratic ethics.

To justify pharmaceutical products, the industry had to document that the drugs were beneficial for patients. The drugs had to be more efficient for healing and cure than traditional psychosocial consciousness-based treatments—now called placebo cures—and the harm they inflicted must be insignificant compared with the healing benefits.

The solution to this problem emerged over a few decades into the standard toxic trial on animals followed by the randomized clinical testing (RCT) on humans used today. In the latter, the potentially beneficial drug is tested double-blind against an inert placebo pill, containing calcium carbonate (chalk), sugar, or a similar, biologically inert drug. Potential effects and suspected adverse effects and events are measured and counted.

Animal testing has during the years proven less effective than expected, as the human body reacts differently toward many drugs than the small mammals we most often use for lab-tests of toxicity. The tests of the drugs on humans often also ignore important poisonous aspects of the drugs, as we have seen many examples of, lately i.e. with the problematic drug Vioxx (2), a seemingly harmless non-steroid pain-killer that greatly increases the risk for acute myocardial infarction and stroke. The medical society has developed a kind of tolerance toward such failure to detect severe and lethal adverse effects of drugs, as it is impossible to test for everything in a pharmaceutical testing. Still, this reminds us all of the importance of choosing a non-drug treatment for our patients, if such a treatment exists at all.

In this paper, we will show that much more fundamental problems exist with the standard RCT procedure used by the pharmaceutical companies to obtain approval and sell their pharmaceutical products than just the occasional failures to detect the toxic effects of drugs. We have analyzed these problems and found them
severe so that we believe RCT to be questioned as a sound basis for evidence-based medicine. At the end of this paper, we therefore suggest more rational methods for testing pharmaceutical drugs.

FOUR FUNDAMENTAL PROBLEMS OF RCT

The RCT is based on four assumptions:
1. The placebo effect is well represented by a placebo pill
2. It is possible to make a double blind test
3. The beneficial and harmful effects of the drugs are fairly measured
4. The timeframe used for the RCT test is reasonable.

The fifth condition for RCTs to be valid is that they are conducted by people without a personal interest in the outcome of the study as such a special interest is likely to induce a bias in the study (3). The problem of bias in pharmaceutical studies based on RCTs is well known and therefore not the subject of this paper. We will discuss these assumptions one by one.

Regarding assumption 1: The placebo effect is well represented by a placebo pill

The core in the classical Hippocratic non-drug treatment is the intimate relationship with the physician—the physician is the tool, so to speak (4). The importance of the close relationships with the physician for the size of the placebo effect has recently been documented in the British Medical Journal (5). The relative importance of talking and touching has been investigated in a recent study by our group (6), and we found the combination of the two principal interventions to be important for inducing a large placebo effect, indicating a strong synergy. The size of the placebo effect has also been established now from several dozens of studies of holistic non-drug medicine done through the last three decades (7,8). We concluded in the two reviews that most health conditions can be treated, and that one patient in two or three is normally cured with the most effective types of placebo-treatments. It has thus been documented that the placebo effect is as tremendously powerful as medicine, even if the patient has a severe heart condition (9,10).

The effect of the placebo-pill used in pharmaceutical RCTs has recently been investigated and the conclusion was that it had no effect at all (11). The authors concluded that the placebo effect (in the RCTs tested) did not exist at all, when compared with no treatment.

Conclusion: The assumption that the placebo effect is well represented with a placebo pill in the RCTs is therefore false.

Regarding assumption 2: It is possible to make a double blind test of a drug vs. (passive) placebo

Most if not all biologically active drugs give an internal clue to the test person that he or she has actually received an active drug. This clue activates a placebo effect called active placebo (12), and if such a drug is tested against normal (passive) placebo, it will induce a placebo effect that by itself will create the result that this is an effective drug. Therefore, to make a test double blind is not possible as the blinding is destroyed by the internal clues of active drugs.

The size of the active placebo effect in psychoactive drugs has recently been established in a Cochrane meta-analysis of antidepressants vs. active placebo (13). The authors found that the effect of antidepressants practically disappeared if tested against active placebo compared with the
normal “passive placebo” pill and when tested against passive placebo, one patient in three was helped.

The latter result is the basis for the marketing of antidepressants today; obviously the conclusion that the anti-depressants help is not justified. We find it possible to extrapolate from this type of drugs to all psychoactive drugs in use today.

The situation is even worse. If you give the patient a poison that gives an internal clue, it will always come out better than the passive placebo used for comparison. The test-substance in the Cochrane meta-analysis was exactly that: drugs that only gave adverse effects and no beneficial effects. So the way the double-blind test is designed favors poisonous drugs for non-poisonous drugs.

In theory, all drugs could be tested using active placebo drugs of similar toxicity, but as this is highly predictable to give the same results as in the above-mentioned Cochrane study, the pharmaceutical industry is not likely to induce this procedure by itself, despite its logical necessity.

Conclusion: The assumption that it is possible to make a double blind test of a drug vs. (passive) placebo in the RCTs is therefore false.

Regarding assumption 3: The beneficial and harmful effects of pharmaceutical drugs are fairly measured

Only with a fair measure of beneficial effects can the effect be evaluated if the drug is useful as a medicine, and only with a fair and similar measurement of benefits and harm is it possible to compare the two for evaluation if the drug all in all is beneficial or harmful to the patient. Many new drugs will reach the clinic or office based on their ability to affect some presumably disease-related measure (i.e., glucose, cholesterol, blood pressure), which are readily measured and can serve as a disease marker. Clearly, something that influences cholesterol but has no effect on cardiovascular disease would not be of much use; however, one would have to learn about this in a stepwise manner. Other strategies for testing drugs focus on local symptoms rather than global states of health and quality of life. Only with global measures can we really know if a pharmaceutical drug of benefit to the patient.

Today we have a number of established and validated global measures of health. We can easily measure self-rated health, self-rated physical health, self-rated mental health, and global quality of life, using small easy-to-use questionnaires (WHOQOL5, QOL1, QOL5, QOL10) (14-16).

We also know that such subjective measures of health are stronger predictors of survival and future health than any objective health measure (17-21). With such measures it is easy to evaluate the total effect of a drug on health and quality of life. This effect can also easily be followed over time. The cost for such testing is minimal and the information gained essential. A global quality of life measure detects the combined effect of benefits and harms from a drug. Many such measures exist, but they are rarely used in RCTs today.

Despite the possibility of using global measures, when they develop and test a drug, pharmaceutical companies most often focus on only one or a few local measures to document positive outcomes. Global benefits in health and quality of life is therefore not tested, nor is global harm. The focus on specific adverse effects and events is further enhancing the drugs chances of looking good in the RCT test, whereas the possible damaging impact on a global scale on health, quality of life, and general level
of performance that would make the drug look very bad is not measured.

When a new drug is marketed, physicians lack this crucial information, and when they do have it, physicians automatically assume that the symptom alleviated by the drug is more important to the patient than the adverse effects induced by the drugs. In this way, drugs that are more harmful than beneficial to health and quality of life can still pass the RCT and come out as a beneficial drug.

Our own analysis of the relative harm and benefits of antipsychotic drugs showed antipsychotic drugs to be about 100 times more harmful than beneficial (22). We recently found a similar situation for cancer chemotherapy (23).

As time goes by, adverse (toxic) effects according to the science of toxicology often tend to accumulate and beneficial effects tend to diminish; it is therefore very important to observe the long term effects of the drugs. In practice, this is almost never done by pharmaceutical companies.

Other problems are that negative results are almost never published, giving a very strong publication bias; when all data are collected in a field, the results are often much more negative than if the industry had just published its positive results, as we saw with the huge meta-analysis of cancer chemotherapy done by Abel (24-27), which concluded that chemotherapy shortens life and destroys the quality of life for almost all types of cancers (the epitheloid cancers).

Conclusion: The assumption that the beneficial and harmful effects of the pharmaceutical drugs are fairly measured in RCTs is therefore false.

Regarding assumption 4: The time-frame used for the RCT test is reasonable
If the active placebo effect of a toxic drug is used in medicine, there will be two phases, a positive phase (the active placebo phase), where the patient feels lifted, motivated, and helped due to the active placebo effect and after this, a negative phase (a toxic phase), where the patient pays the price of being helped by a toxic drug.

The ideal use of a toxic drug—like strychnine that was used by allopathic physicians around the 1900s—was a short, strong intervention. If the treatment period was too long, then the immediate benefit would be destroyed by the harm caused in the long run. We therefore know that an RCT-test involving a strong element of active placebo from the toxic effects of biologically active chemicals must be thoroughly tested for the whole period of time that it is being used by patients to monitor the total effect on the patients.

In our meta-analysis of antipsychotic drugs (22), we learned that a positive effect found in a short term measurement at 6 months often is reduced to half the effect after 12 months, and we presume that this tendency continues though time, making it mandatory to test positive effects for 2 years or more, as many patients are treated with the drugs for years, in the belief that a short-term effect is also preserved in the long term.

In the same way, the adverse effects (side effects) and adverse events (negative events) tend to accumulate through time. As an example, schizophrenic patients more and more frequently commit suicide as treatment with antipsychotic drugs continues. A few percent of patients take their own life in the beginning of pharmaceutical treatment (28), with this fraction growing to 15% as time goes by (29).

Swedish researchers have suggested that suicide is caused by drug-induced depression (30). The reason for the increased rate of patient suicide though time could very well be a more and more severe depression
induced by the accumulated toxic effect of the anti-psychotic.

Psychiatric patients treated with pharmaceutical drugs also have a higher tendency to die spontaneously (31), presumably because of accumulated toxic effects. It is therefore of extreme importance to continue the measuring of toxic adverse effects and adverse events during the long term (2 years or more, or for 5-20 years if patients take the drugs often for so long).

If the appropriate time frame for RCTs is not used, then the whole test becomes meaningless. Unfortunately most pharmaceutical companies only test their product over a short term, often only for three months. This approach seems to be a strategy to hide the adverse effects of the drugs, which is unacceptable.

**Conclusion:** The assumption that the timeframe used for the RCT test is reasonable is therefore false.

**COMBINED EFFECT OF ALL FOUR ERRORS**

The first error, not testing pharmaceuticals against the traditional psychosocial intervention that holistic physicians have been doing for millennia is giving the pharmaceutical industry an easy way out or no competition at all. Basically all drugs can win this race.

The second error is changing the sign of the test from plus to minus—toxic drugs are perceived as beneficial drugs due to the active placebo effect. This approach is problematic and makes the present RCT-procedure misleading.

The third effort, local non-global testing (a local symptom or a disease marker) ignores the possible, negative global effects on the patient’s health caused by toxic effects of the drugs. The focus on local effects separates positive effects from adverse effects, making it possible to ignore that the harmful effects are stronger than the beneficial and allows the industry to conclude that the drug has beneficial qualities for specific symptoms. Because of this way of testing, even a very toxic drug can pass the RCT test and come out as beneficial.

The fourth effect, to test only in a short term period, is boosting the positive effects caused by the active placebo effect and hiding the true, adverse effects of a drug used in the long run.

By combining all four errors, the pharmaceutical industry has managed to set up a RCT procedure that can make almost any drug look like a beneficial pharmaceutical medicine with only modest harm done to patients.

In the documentation, drugs will appear as clinically beneficial drugs with clinically less significant adverse effects and events. Here we have a situation that is clearly not acceptable. The RCT procedure needs instant revision and should not be used in its present form for future clinical testing of pharmaceuticals. All drugs tested with the RCT procedure have to be retested as we cannot rely on the results of the present RCT-test procedure.

Without a doubt, all four errors have been used individually because of their ability to improve the way that pharmaceutical drugs come out of RCTs. We doubt that the highly problematic combined effect—that toxic drugs are made to look like beneficial medicine—is made intentionally as the effect of the pattern of the four errors combined is somewhat difficult to understand.

On the other hand, there have been times when pharmaceutical drugs subjected to large meta-analyses have turned out to be only harmful and not beneficial at all, as we have seen (13,24). Another example is anti-
psychotic drugs, for which Adams et al (32) found in a large Cochrane metaanalysis that these drugs do not improve mental health (“mental state”) at all, with the drugs having many very common and severe adverse effects. Adams et al (32) also found that the new generation of antipsychotic drugs is not more beneficial or less harmful than the original drug, Chlorpromazine, despite that the industrial RCT tests of the new generation drugs often show an improvement.

The pharmaceutical industry has had both time and plenty of occasions to reflect upon the contrast between the results of the single RCT-based study made by the industry and the conclusions of the large meta-analyses made by independent researchers.

We believe that the pharmaceutical industry has done its own critical analyses, very similar to the one we are presenting here, but has not taken the consequences and changed the RCT procedure. This is in part because scientific journals are accepting the RCT procedure as it is and partly because it is good business for the industry.

**HOW SHOULD BIOMEDICAL INTERVENTIONS BE TESTED?**

If one wants to keep the design of the RCT, then one should use an active placebo of sufficient strength, global outcomes, and sufficient observation time. We have an ethical problem with the use of active placebo drugs as they must be as toxic as the drug we are testing, but without positive medical qualities. It is not simple to distribute toxics to thousands of innocent control patients.

For chronic patients, a simple schedule must be preferred: Simply treat chronically ill patients—patients that have not been better for years—and see if they improve on some global level—health, quality of life, or performance. Follow them for a few years and see if the induced improvement is permanent. Use NNT (number needed to treat) and NNH (number needed to harm) numbers to express the effectiveness and use, if possible, at all outcomes “cured or not cured” in combination with self-rated physical health, self-rated mental health, and self-rated quality of life.

If a pharmaceutical treatment cures a fair fraction of the patients, say one in 2, 3, or 4, and does not have significant adverse effects, then this is a valuable drug. If not, if it cures only one in 50, and if it has significant adverse effects, then the drug is of no medical value. If there is a more effective, or similarly effective, non-drug treatment, then the pharmaceutical treatment is of no value as there will always be some adverse effects from drugs.

This procedure of curing chronically ill patients and using them as their own control is simple and efficient, and can be used with all types of chronic patients (33). The randomization to no treatment is less valuable, as most of these patients will go to some kind of CAM treatment if not treated medically. If classical, Hippocratic holistic medicine is used as a control in the study and we recommend that the research follows the open source protocol for clinical holistic medicine (34).

For acute patients, randomization is still necessary. The most logical thing to do is to randomize to holistic medical treatment; there are many small units with holistic physicians, who have documented their efficacy. If a holistic medical treatment unit is not available for a specific disease, then it will be necessary to train a group of physicians to do it, or if this cannot be done, randomize to no treatment. When a patient receives no treatment from a doctor, the patient—i. e. in an acute psychotic
crisis—is likely to assume more responsibility for his or her own life, and this itself has a strong curative effect.

We understand that a serious proposal to create an NGO (non-government organization) to reevaluate every approved drug on the market would involve a breathtaking commitment of resources, but it can and should be done. We estimate that a research hospital specifically established for comparing biomedicine and classical holistic medicine would cost around or 150 million EURO or $US 200 million to establish. This amount is still not much on an industrial or national scale.

DISCUSSION

Most researchers acknowledge that there is no risk-free ride when a patient takes a drug to obtain a benefit; every drug has some adverse effects. We have found that the way that the RTC tests the medical value of pharmaceutical drugs today tends to create the impression that a drug that has no beneficial effect at all, but only harmful adverse effects, can still appear as an effective, useful medicine. Toxic drugs tested with the RTC method can thus be sold as medicine.

The RCT-procedure is built on false assumptions in our opinion and has strong built-in bias in favor of the drugs. We know that biologically active drugs can be toxic and it is therefore of extreme importance that we are able to make a fair test of pharmaceutical drugs to ensure benefit to the patient and do no harm. The problems relate to the choice of placebo types, to the outcomes used, and to the observation times, not to mention all the other types of well-known bias like the withdrawal of negative results from publication, which could explain the findings of Abel (25).

We basically see all this as a political and commercial problem rather than as a scientific problem. The scientific problems of RCTs can easily be solved: In principle, the CRT can test the benefit and harm of a drug using randomized, double-blind testing compared with an active placebo using the drugs we have suggested and global outcome measures. This approach could easily be accomplished without any technical or scientific problems.

But the industry tests its pharmaceutical drugs in such a way as to optimize the appearance of the drugs, which is only logical from a commercial perspective. The pharmaceutical companies make the drugs look as beneficial and as harmless as they possibly can. It is important to recall that the way in which the drugs are tested has been created by the pharmaceutical industry. In our opinion, these drugs have been uncritically approved both by responsible government institutions and by physicians.

Academic institutions have in general also approved the standard RCT method for pharmaceutical drug testing without being critical enough in our opinion. We are now in a difficult situation because drugs have been accepted, but they might be harmful and not beneficial to patient health and quality of life.

We know that about 50% of citizens in countries with socialized, free biomedicine are chronically ill (35). Analyses have shown that only a small fraction of these patients are helped by drugs (36). The deteriorating health of the population might be explained directly from the toxic effect of the many pharmaceutical drugs given to the population.

The solution to the difficult situation is to test all drugs on the market again. Pharmaceutical drugs must be tested by an organization that does not have commercial interests in the drugs. Such organizations are hard to find and might have to be
created from the bottom, finding researchers without personal interests in medicine. Such an organization must preferably be an NGO as strong lobbyism from the pharmaceutical industry continues to plague the public health care system.

The possible result of such testing could very well be that classical holistic medicine inducing healing of mind and body—often called salutogenesis—may be found preferable to symptom-blocking drugs, which do not heal the person (36-40). A broader application of subjective health and quality of life measures would constructively impact the RCT test.

CONCLUSIONS
The standard RCT testing of pharmaceutical drugs in double blind trials as compared with placebo has so many problems that these sum up to a fatal error: a drug with only toxic qualities is likely to appear as a beneficial medicine. The primary single cause for this is that toxic drugs always have an active placebo effect that makes the drug look beneficial in the RCT-test. This cast serious doubt that the RCT-procedure in its present form is scientifically valid.

The way that clinical outcomes are chosen in the tests—with a focus on local symptoms or disease markers instead of global states—makes it further impossible to compare positive and negative effects.

Finally, the short time frame of testing makes the positive active placebo effect dominate over the negative pharmacological drug effect of a toxic drug. Therefore, when the standard RTC is used, a toxic drug with no beneficial pharmacological effect is likely to be approved as a beneficial pharmacological medicine.

We conclude that effects of drugs documented with the standard RCT-test procedure used by the pharmaceutical industry today are not “evidence based”. As a consequence, we cannot exclude the possibility that some of the pharmaceutical drugs in use today are likely more harmful than beneficial, despite being documented as primarily beneficial. We therefore have to re-test all pharmaceutical drugs documented with the RCT-test. This can be done using the simple test on chronic patients with randomization against no treatment, or better against the traditional placebo cure by classical holistic medicine, or in acute medicine, using a randomized test using active placebo, global outcomes, and sufficiently long test times.

Testing must be done by individuals and organizations having no personal, commercial, or political interest in medicine. We strongly advise that NGOs be empowered to do the testing, as all governmental organizations are strongly influenced by the lobby of the pharmaceutical industry.

We recommend the establishment of a research hospital dedicated to the testing of medicine that could compare the effects of pharmaceutical drugs with the effects of classical holistic medicine, the original placebo cure, for each clinical condition. We estimate that this could be done for about $US 200 million or 150 million EURO, which is not much on a national or industrial scale.

ACKNOWLEDGMENTS
The Danish Quality of Life Survey, Quality of Life Research Center and the Research Clinic for Holistic Medicine, Copenhagen, was from 1987 till today supported by grants from the 1991 Pharmacy Foundation, the Goodwill-fonden, the JL-Foundation, E Danielsen and Wife’s Foundation, Emmerick Meyer’s Trust, the Frimodt-Heineken Foundation, the Hede Nielsen Family Foundation, Petrus Andersens Fond,
Wholesaler CP Frederiksens Study Trust, Else and Mogens Wedell-Wedellsborg’s Foundation and IMK Almene Fond. The research in quality of life and scientific complementary and holistic medicine was approved by the Copenhagen Scientific Ethical Committee under the numbers (KF) V. 100.1762-90, (KF) V. 100.2123/91, (KF) V. 01-502/93, (KF) V. 01-026/97, (KF) V. 01-162/97, (KF) V. 01-198/97, and further correspondence. We declare no conflicts of interest.

REFERENCES
8. Submitted to Soc Indicators Res.
30. SBU-rapport nr. 133/1 og 133/2. Behandling med neuroleptika. Stockholm: Statens beredning för utvärdering av medicinsk metodik, 1997;2. [Swedish]
33. Ventegodt S, Andersen NJ, Merrick J. The square curve paradigm for research


