J. Pharm. Pharmacol. 1999, 51: 237–238
Received October 29, 1998
Accepted November 2, 1998

Guest Editorial – The History of the Double Blind Test and the Placebo

FREDERICK C. STRONG III

State University of Campinas, Campinas, SP, Brazil and Bucknell University, Lewisburg, PA 17837, USA

In the course of some personal research in the area of food intolerance, I found it essential to undertake a double blind test. As I learned more about the concept, I became convinced that it is one of the most important tools that has been developed this century for the experimentalist in pharmacology and allergy studies. Its history deserves to be documented and, likewise, that of the placebo.

The *Oxford English Dictionary* gives the following definition of double (-) blind:

"Applied to a test or experiment conducted by one person on another in which information about the test that may lead to bias in the results is concealed from both the tester and the subject until after the test is made; orig. used of tests for determining the efficacy of drugs."

It then gives references to three publications that incorporated the double blind test prior to a review in 1954 and the subsequent explosion in its use described in another reference. These provided a starting point. I then succeeded in finding five more papers that used the double blind test before 1954.

The first published account of taking steps to avoid bias in interpreting experimental results by deliberately withholding information from the experimenter and the patient appears to be that by Gold et al (1937) in the *Journal of the American Medical Association*. It was a study of the application of theobromine and aminophylline to the treatment of cardiac pain:

"Various devices were employed to guard against directing the patient's judgment... In a further attempt to eliminate the possibility of bias, the questioner usually refrained from informing himself as to the agent that had been issued until after the patient's appraisal of the period had been obtained... The procedure was based on this general formulation; namely, if the relief of pain during the use of the xanthine is due to the specific action of the drug, the patient should be able to distinguish its effects, and to do so repeatedly, from the effects of a placebo given under similar conditions and in such form as to preclude its detection by the patient through any means other than the relief of pain... Summary. The data consisted of the patients' judgements regarding changes in pain. These data were secured in a manner relatively free of bias by the use of the 'blind' test."

The authors were Harry Gold, Nathaniel T. Kwit and Harold Otto. This report and seven subsequent research reports are described as being from the Department of Pharmacology of Cornell University Medical College, Beth Israel Hospital or the Hospital for Joint Diseases, or from a combination of the establishments, all in New York City. According to Roueché (1960), their blind test was based on work done by the British experimental psychologist H. H. R. Rivers in 1906.

After an interval of five years, Gold et al (1942) published a paper on the bioassay of digitalis in humans by means of electrocardiograms, using the "effect on the RT-T group by the 'blind test'...". Six years later, in another paper testing the use of aminophylline for angina patients, Bakst et al (1948) reviewed the 'blind' technique of Gold et al (1937) and then adopted it.

The following year, Travell et al (1949) published a study of the effect of α-tocopherol on chest pain in cardiac patients:

"Blind test methods were applied both in the collection and interpretation of data. One person issued the medication and never examined the patients. Thus, none of the examiners knew which patients were receiving the vitamin and which ones the placebo. Even judgements regarding the final result were made in each case without knowledge of the nature of the material administered."

The first four papers spoke of the "blind test". The next (Greiner et al 1950), on the value of khellin for cardiac pain relief and published in the *American Journal of Medicine*, was the first to use the term "double blind" (incidentally, the fourteen authors wrote double blind without a hyphen).

In a study "conducted according to stringent blind-test methods", and published the same year, Rinzler et al (1950) investigated the failure of α-tocopherol to influence chest pain in patients with heart disease.

With regard to the bioassay of diuretic agents, local reactions and systemic toxicity, Greiner et al (1951) wrote:

"Here symptoms and judgement are involved and for that reason, the 'blind test' is imperative, especially in case of local reactions (pain, burning, redness, swelling, induration). The 'blind test' may be applied by using solutions from containers labeled with consecutive numbers, the identity of the compound remaining unknown to those making the injections and eliciting the patient's history."

A final example is Rinzler et al (1953) who described the control of their procedure on the effect of heparin in effort angina as follows:

"To eliminate the factor of unconscious bias the double blindfold method was employed throughout. This meant that the study was conducted by a team, and that not merely the patients but also the physicians who questioned and examined them, injected the solutions and later assessed the data were unaware of the nature of the coded solution given to any particular patient."

The above research paper is the eighth cited here, published between 1937 and 1954, that employed the double blind procedure. There may have been a few more, but probably not many. It is interesting to note that only the third paper (Bakst et al 1948) did not have an author who had not previously been a co-author of a paper utilizing the double blind test.

In a review lecture on clinical pharmacology, Gaddum (1954), at the University of Edinburgh, stated:

"Errors of assessment may occur when the results of the treatment are assessed by someone who may be too hopeful or too sceptical... One method of avoiding this type of error is to depend entirely on objective measurements of such things as temperature or weight. This is seldom very satisfactory, and it seems a pity to pay no attention at all to the doctor's opinion or to the patient's opinion. The only safe way to obtain unbiased opinions from either of them is to make them express their opinions without knowing whether the patient received an active drug or not. This is known in America as a double blind test."

Following Gaddum's review, use of the double blind procedure grew and grew. A long article on the placebo and the first double blind test appeared in the *New Yorker* (Roueché 1960). An author in the *Lancet* (Anon. 1961) commented:

"Statistics and certain concepts, such as double blind trial, are on everyone's mind today..."

Any history of the double blind test should include the history of the placebo since it is an integral part of the procedure. Gaddum (1954) was not in agreement with the use of the term placebo. Referring to the double blind test, he stated:

"It generally involves giving the controls dummy treatment which cannot be distinguished from the real treatment. If the treatment is given in the form of tablets then the controls receive dummy tablets indistinguishable in appearance, taste and smell from the real tablets... Such tablets are sometimes called placebos, but it is better to call them dummies. According to the *Shorter Oxford Dictionary*, the word placebo has been used since 1811 to mean a medicine given more to please than to benefit the patient."

Besides describing placebos as materials given to patients to please them, the *Oxford English Dictionary* even cites references to physicians "writing prescriptions for placebos". However, Gaddum's plea for the use of dummy rather than placebo appears to have fallen on deaf ears. Placebo is generally understood today to mean an inert substitute for a real drug or food sample.

On December 11, 1998 an article appeared in the *Wall Street Journal* that was on the use of the double blind procedure in surgery. Johannes (1998) reported that a test using a fake procedure of surgery for the pain of angina was attempted in the 1950's and provoked a backlash. However, she reported: "But things are changing... Sham surgeries have been used in at least five recent or current trials of therapies: three for Parkinson's, one for cancer pain and one involving the knee. In some cases it is impossible to keep the surgeons in the dark about who is getting the placebo and who isn't."

Acknowledgements

The author is grateful for the assistance of Patricia Housner, James Van Fleet, Georgia Smith, Lynda Thaler and Kay Knapp of Bucknell University's Bertrand Library, and Nathaniel T. Kwit Jr.

References

Anon. (1961) The pharmaceutical persuaders; the industry, the doctor, and the clinical trial. Lancet i: 421–423

Bakst, H., Kissin, M., Liebowitz, S., Rinzler, S. (1948) The effect of intravenous aminophylline on the capacity for effort without pain in patients with angina of effort. Am. Heart J. 36: 527–534

Gaddum, J. H. (1954) Clinical pharmacology. Proc. R. Soc. Med. 47: 195–204

Gold, H., Kwit, N. T., Otto, H. (1937) The xanthines (theobromine and aminophylline) in the treatment of cardiac pain. J. Am. Med. Assoc. 108: 2173–2179

Gold, H., Cattell, McK., Otto, H., Kwit, N. T., Kramer, M. L. (1942) A method for the bio-assay of digitalis in humans. J. Pharmacol. Exp. Ther. 75: 196–206

Greiner, T., Gold, H., Cattell, McK., Travell, J., Bakst, H., Rinzler, S. H., Benjamin, Z. H., Warshaw, L. J., Bobb, A. L., Kwit, N. T., Modell, W., Rothendler, H. H., Messeloff, C. R., Kramer, M. L. (1950) A method for the evaluation of the effects of drugs on cardiac pain in patients with angina of effort. Am. J. Med. 9: 143–155

Greiner, T., Gold, H., Bliss, C. I., Gluck, J., Marsh, R., Mathes, S. B., Modell, W., Otto, H., Kwit, N. T., Warshaw, L. (1951) Bioassay of diuretic agents in patients with congestive failure. J. Pharmacol. Exp. Ther. 103: 431–440

Johannes, L. (1998) Sham surgery is used to test effectiveness of novel operations. Wall St. J. Dec 11, 1998; 232 (115): A1, A8

Rinzler, S. H., Bakst, H., Benjamin, Z. H., Bobb, A. L., Travell, J. (1950) Failure of alpha tocopherol to influence chest pain in patients with heart disease. Circulation 1: 288–293

Rinzler, S. H., Travell, J., Bakst, H., Benjamin, Z. H. Rosenthal, R. L., Rosenfeld, S., Hirsch, B. B. (1953) Effect of heparin in effort angina. Am. J. Med. 14: 438–447

Roueché, B. (1960) Annals of medicine – placebo. New Yorker 36 (Oct. 15): 85–103

Travell, J., Rinzler, S. H., Bakst, H., Benjamin, Z. H., Bobb, A. L. (1949) Comparison of effects of alpha-tocopherol and a matching placebo on chest pain in patients with heart disease. Ann. N. Y. Acad. Sci. 52: 345–353