

Department of medical history

Powerful placebo: the dark side of the randomised controlled trial

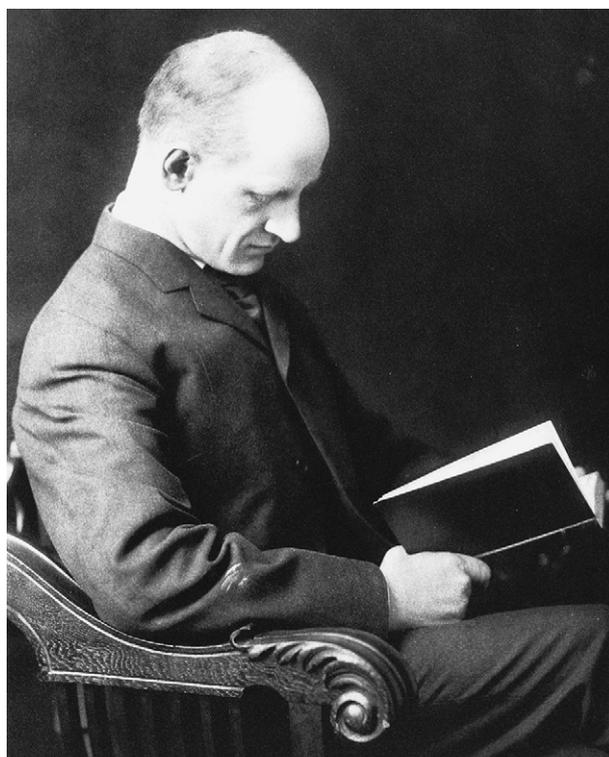
Ted J Kaptchuk

The placebo went through a dramatic metamorphosis in the years after World War II as the double-blinded randomised controlled trial (RCT) developed. Until 1945 the placebo was a “morally” useful but innocuous management tool without curative or symptomatic consequences. By the time the double-blind randomised controlled trial took form and began to establish itself, around 1955, the placebo was imbued with powerful therapeutic effects and its ethical clinical use was more generally being questioned. In 10 years, the placebo changed from what was called the “humble humbug” to an entity with occult-like powers that could mimic potent drugs. It may be that efforts to bring the precision of science into the evaluation of efficacy with the RCT has its own form of confusion and darkness.

Pre-RCT placebo

When pre-World War II paternalistic ethics prevailed, informed consent was not a standard of care.¹ Physicians were generally comfortable with benevolent deception and a “polychromatic assortment of sugar pills” was routinely swallowed by patients.^{2,3} In 1903, Richard Cabot (1868–1939), the eminent professor at Harvard Medical School, described the placebo’s persuasiveness. He was “brought up, as I suppose every physician is, to use placebo, bread pills, water subcutaneously, and other devices . . . How frequently such methods are used varies a great deal I suppose with individual practitioners, but I doubt if there is a physician in this room who has not used them and used them pretty often . . . I used to give them by the bushels”.⁴ Cabot’s remarks are confirmed in many other observations of medical practice. For example, in 1807, Thomas Jefferson (1743–1826) penned a description of what he called the “pious fraud” and noted that, “one of the most successful physicians I have ever known has assured me that he used more bread pills, drops of coloured water, and powders of hickory ashes, than of all other medicines put together”.⁵

Peculiarly (at least in the current view), the bread pill was generally thought to have no consequences other than giving patients (especially ignorant and complaining ones) peace of mind. A medical dictionary published in 1785 described the placebo as “calculated to amuse for a time, rather than for any other purpose”; a dictionary from 1811 depicted it as “given more to please than to benefit the patient”; and until the 1950s medical dictionaries echoed this definition of an inactive and innocent management contrivance.⁶ The main objections to this prevailing view can be found in the few sympathetic discussions of such unorthodox practices as



Richard Clarke Cabot

hypnosis, where the power of the imagination was accepted.⁷ If there was a “medical” value for a placebo it was as an occasional diagnostic tool to separate imaginary “psychological” symptoms from real problems.⁸ In 1945, an influential article stated that the placebo was a valuable device “to smooth [the patient’s] path”, which “cannot harm and may comfort” patients, especially the “ignorant . . . disappointed and displeased . . . hopeless, [and] incurable case[s]”.⁹ A 1954 *Lancet* article, entitled “The Humble Humbug”, depicted the swan song on this old-fashioned understanding of the placebo; “a means of reinforcing a patient’s confidence in his recovery, when the diagnosis is undoubted and no more effective treatment is possible; that for some unintelligent or inadequate patients life is made easier by a bottle of medicine to comfort their ego; that to refuse a placebo to a dying incurable patient may be simply cruel; and that to decline to humour an elderly ‘chronic’ brought up on the bottle is hardly within the bounds of possibility”.¹⁰

RCT placebo effect

In 1955, the modern biomedical concept of the placebo makes its first definitive appearance with the publication of Henry Beecher’s (1904–1976), “The Powerful Placebo” in the *Journal of the American Medical Association (JAMA)*¹¹ Beecher, a distinguished medical researcher at

Lancet 1998; **351**: 1722–25

Center for Alternative Medicine Research, Beth Israel Deaconess Medical Center, Harvard Medical School, Boston, Massachusetts 02215, USA (T J Kaptchuk omd)

Harvard Medical School, summarised and mathematically presented a perspective that had been developing in a few elite biomedical research centres since 1946.¹² Beecher used a proto-meta-analytic method to aggregate the percentage of patients satisfactorily relieved by a placebo across 15 clinical trials. Of 1082 patients, 35.2 (SD 2.2%) experienced therapeutic benefits. This number, which Beecher called “average significant effectiveness” did not measure the exact magnitude of improvement, although that is how most people have subsequently interpreted his paper. Beecher argued forcefully that placebos alleviated beyond psychological problems by citing evidence that “these powerful placebo effects . . . can produce gross physical change”, which “include objective changes at the end organ which may exceed those attributable to potent pharmacological action”.¹¹ In an important revision of old orthodoxy, Beecher considered this placebo effect to operate regardless of the intelligence of a person.

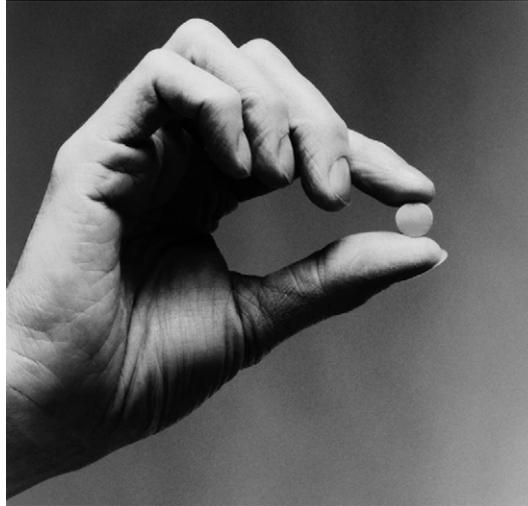
Beecher explicitly assumed an additive model of placebo effects, “The placebo effect of active drugs is masked by their active effects . . . The total ‘drug’ effect is equal to its ‘active’ effect plus its placebo effect” (quotes in the original).¹¹ This premise took for granted that the active drug response results partly from a placebo effect and that the placebo effect buried in the active arm is identical to the placebo effect of the dummy treatment.¹³ The placebo was a single and stable “power” that behaved in a consistent manner. The new placebo effect, the oxymoron-like enigma of an effect produced by something that is inert came to haunt biomedicine. Still to this day, extensive examination leaves scientists and philosophers to conclude that “the placebo concept as presently used cannot be defined in a logically consistent way and leads to contradictions”.¹⁴ Conflicting explanations—expectation, faith, classic conditioning, anxiety relief, symbolic processes, patient-doctor relationship, self-perceptions—vie for acceptance. The placebo effect has attributes of a neo-mesmeric energy. It is radically protean: placebo responders seem to react or not capriciously,¹⁵ and for the same disease the placebo effect can vary between 0% to 100%, when compared with identical drugs.^{16,17} The theurgic placebo can even be more effective^{18,19} or, as noted by Beecher, completely reverse the outcomes of known powerful pharmacological drugs.²⁰

What happened in the 10 years between 1945 and 1955, when the placebo shifted from insipid decoy to a mischievous genie that could trick the most discerning clinician?

RCT and creation of the modern placebo

World War II has been called the great divide in medical research and certainly, for the placebo, this is true. In fact, the “powerful placebo” was born in the vortex of one of medicine’s most momentous transitions. Before

World War II the evaluation of new therapeutics was made by recognised leaders of the medical profession whose judgments were based on clinical impressions, and on rare occasions, poorly controlled evidence.²¹ In an effort to impose an objective and scientific discipline onto the extraordinary postwar expansion of medical research, the components of the double-blind RCT were adopted and coalesced in the years after the war.²² The major features of these innovations included blind assessment (usually meaning a placebo control), random assignment to comparable groups, and inferential statistics as a surrogate for determinism.²³



A green placebo?

The postwar placebo effect resulted from an almost sleight-of-hand shift in the placebo’s operational meaning in the new RCT model. Instead of an inert sham given to individual patients, the placebo became the emblem for all the healing occurring in the disguised “no-treatment” arm of an RCT. The “placebo effect” encompassed all “nonspecific effects” that did not depend on the treatment in the active arm. The “powerful placebo” became a hodge-podge of non-linear, difficult to

quantify, remnants collected under the rubric of the dummy control of an RCT. Anything that threatened the fastidious detection of a predictable cause and effect outcome was conveniently disposed of in a repository labelled the “placebo effect”. This new concept of placebo was much larger both in meaning and power than its predecessor. It incorporated many contributors to health outcomes such as natural history, routine medical and nursing care, and the “art” of medicine that had once been clearly distinct from the deception of an inert pill.

Powerful placebo and acceptance of RCT

Medical proponents of the RCT were under pressure to convince their colleagues of the RCT’s value.²³ Few physicians wanted randomly to assign treatment to patients, forgo the individualisation of therapy, and withhold promising new therapies. Austin Bradford Hill (1897–1991), the designer of the first randomised trial, many years later confessed that he “deliberately left out the words ‘randomisation’ and ‘random sampling numbers’ at the time because . . . I might have scared them [collaborating physicians] off”.²⁴ Physicians resisted treating patients as so “many bricks in a column” and the “elimination of the responsibility of the doctor to get the individual back to health”.²⁵ In the same issue of *JAMA* as Beecher’s paper, another research team concluded, “we seriously doubted whether the double blindfold technique was a valid method of distinguishing between the effectiveness of analgesic agents”.²⁶

For elite researchers, the moral imperative for scientific method in therapeutic evaluation was critical. They realised that the imprecision of standard methods was a hazard for the health of nations. (The general public was mostly unaware of any problem until the much later thalidomide tragedy). For these reformers, “the powerful placebo effect” became a major argument used to

persuade the medical profession to accept the placebo-controlled RCT. The greater the placebo's power the more necessity there was for the masked RCT itself. Because of this ominous threat, which could distort the judgment of even the most unbiased and conscientious researcher, one needed to adopt the scientific device of placebo control which alone allowed a separation of real from false effects. An enhanced "placebo effect" came to serve a valuable scientific and rhetorical function of persuading colleagues of the necessity of the RCT.

A need to show the placebo's power was understandable. Understanding the phenomena itself was of much less consequence. How the active treatment worked was seen as important. The dummy side of a trial received inadequate attention, poor methodology, and a priori assumptions. The archetypal example of this is Beecher's original study with its many flaws, most of which have gone unnoticed until recently.^{27,28} For example, Beecher did not mention, although he undoubtedly knew, that much of what was labelled a powerful placebo effect was actually regression to the mean, natural history, or concurrent interventions.²⁷⁻²⁹ There had already been some well-designed experiments with two controls (a placebo arm and a "no-treatment no-placebo" arm) which demonstrated no placebo effect.³⁰ These were not included in the study. Beecher retrospectively interpreted the two pre-World War II trials used in his study as having a placebo effect when the original investigators have confidently reported the results observed in the sham arm as being due to spontaneous recovery³¹ or spontaneous variations.³² Also, in the calculation of the quantitative effect of the placebo, Beecher deliberately did not include in his calculation the numbers of patients who got worse with placebo, although this effect was reported in several trials Beecher analysed.³³ Inclusion of this group would have dissolved significant amounts of the placebo effect into equally distributed normal variability.²⁷ In fact, the entire point of Beecher's exercise to establish the "powerful placebo" was to demonstrate persuasively "that 'clinical impression' is hardly a dependable source of information without the essential safeguards of the double unknown technique, the use of placebos also as unknowns, randomisation of administration . . . and mathematical validation of any supposed differences".¹¹ Accurate portrayal of the placebo effect was of less importance than invoking it as a threat to scientific evaluation, the elimination of which would be accomplished by the double-blind RCT. The "new" placebo became both the *raison d'être* for, and the sacrificial victim of, the masked RCT.

RCT and placebo: the light and dark of a partnership

Elite medical reformers created a symbiosis between the RCT and the powerful placebo effect. A new gold standard was constructed to fit the new technical procedures. Until the RCT, medical therapy became legitimate because of beneficial outcomes; after the RCT, a medical intervention was only scientifically acceptable if it was superior to placebo.³⁴ No longer was it sufficient for a therapy to work: it had to be better than placebo. For the first time in history (outside of religious healing scuffles), method became more important than outcome.³⁴ In a self-authenticating manner, the double-

blind RCT became the instrument to prove its own self-created value system. This shift from emphasising outcomes to the purity of the means directly parallels developments in medical ethics where "informed consent" replaced "beneficence" as the pinnacle of the value system.

Presuppositions embedded in the new concept of placebo also helped implement new methods of using frequency statistics to make causal inferences. The critical assumption here was that the placebo effect was a monolithic effect which was present to the same degree and same direction in both the treatment and dummy arms. (Anomalies such as placebo with a larger effect than the real drug or a placebo that could reverse pharmacological activity, were conveniently overlooked as was the possibility of verum and placebo being differentially effected by the context of the RCT or of interacting.) For the emerging RCT model, the treatment and dummy arm of trials were assumed to receive equal and independent amounts of this force; one could simply subtract the amount of placebo effect to determine the presence (or absence) of specific drug effect. The possibility that the placebo effect could act differentially in the two arms was discounted. This "assumption of additivity . . . enable[d] one to infer that the variabilities within the treatments should have a . . . [random] distribution."³⁵ Without the premise of a single placebo effect, commonly used statistical procedures would be confounded.

The placebo had value only as a comparison marker, the magnitude of its absolute power has been an incidental question. Studies attempting mathematically to quantify across trials the magnitude of the placebo effect, like Beecher's, can be counted almost on one hand.^{17,27,36-39} When their results conflict, as they do, they have been ignored or tolerated. Beecher's vintage numbers are still routinely misquoted; they are preferable to any challenge of their written assumptions.

RCT and powerful placebo at 50 years

The powerful placebo is a modern entity constructed in the shadow of the RCT. In the current RCT era, a legitimate therapy must demonstrate an effect greater than a decoy disguised as a real intervention. Yet, under the rhetorical label of powerful placebo lies many rich contributions to health care. These include: nature taking its course; regression to the mean; routine medical and nursing care; regimens such as rest, diet, exercise, and relaxation; easing of anxiety by diagnosis and treatment; the patient-doctor relationship; classic conditioning and learnt behaviours; the expectation of relief and the imagination; and the will and belief of both patient and practitioner.

The placebo effect also includes another often-overlooked consequence of research activity. It is modified by consequences due to the context of the RCT itself.⁴⁰ Issues such as the method of recruiting patients, manner of giving informed consent, procedures for blinding, vehicle of delivery (colour of pills, pills *vs* injection), provider characteristics, provider verbal attitudes, and physical setting of the environment have been insufficiently studied.

But each of the components of the placebo effect has consequences in healing and may have a differential impact on each arm of an RCT. All the variables can

create countless variations in outcomes that have undoubtedly contributed to the haphazard trail that the placebo effect has traced. It is therefore essential, whenever ethically and financially feasible, that research in medicine begins to disentangle the “non-specific” and “art-of-medicine” aspects of healing and therapeutic evaluation. Besides comparing a real intervention and placebo, the inclusion of a third no-treatment no-placebo arm would be helpful to distinguish a perception of a “placebo” effect from the ordinary natural history of a condition.⁴¹ Each “non-specific” effect needs to be disentangled, carefully varied, and systematically studied under controlled conditions.⁴⁰ Preliminary evidence concerning the examination of non-specific effects and contextual effects of an RCT has been valuable, provocative, and contributed to a more refined understanding of the internal and external validity of trials.⁴²

Obviously, the double-blind RCT has meant a tremendous improvement in research and subsequent medical care. In the beginning, the RCT needed a simplistic neo-mesmeric placebo as a looming threat. It is undoubtedly time that the “powerful placebo” be examined in all its myriad facets, otherwise medicine will always have a limited perception of healing. At age 50, the RCT is ready to go through a mid-life crisis and face its dark side.

Acknowledgments

Research for this paper was partly supported by grants from the John E Fetzer Institute, NIH grant U24 AR3441, and the Kenneth J Germeshausen Foundation. The author also wishes to thank David Eisenberg, Howard Brody, Daniel Moerman, Adriane Fugh-Berman, Herman Engelbart, and David Stone for advice, discussion and feedback, and thank Marcia Rich and Janet Walzer for editorial assistance.

References

- Faden RR, Beauchamp TL. A history and theory of informed consent. Oxford: Oxford University Press, 1986.
- Anon. Placebos. *Med Record* 1885; **27**: 576–577.
- Rosenberg C. The practice of medicine in New York a century ago. *Bull Hist Med* 1967; **41**: 223–53.
- Cabot RC. The use of truth and falsehood in medicine: an experimental study. *Am Med* 1903; **5**: 344–49.
- Jefferson T. The writings of Thomas Jefferson. v. 9 PL Ford (ed). New York: G P Putnam's, 1898.
- Shapiro AK. Semantics of the placebo. *Psych Quart* 1968; **42**: 653–95.
- Tuke DH. Illustrations of the influence of the mind upon the body in health and disease designed to elucidate the action of the imagination. Philadelphia: Henry C Lea's, 1884 [1872].
- Shideman DE, Beckman H. Comments on treatment. *Wisconsin Med J* 1958; **57**: 456.
- Pepper OHP. A note on the placebo. *Am J Pharm* 1945; **117**: 409–12.
- Anon. The humble humbug. *Lancet* 1954; ii: 321.
- Beecher HK. The powerful placebo. *J Am Med Assoc* 1955; **159**: 171602–1606.
- Conference on therapy. The use of placebos in therapy. *N Y J Med* 1946; **17**: 722–27.
- Spilker B. Guide to clinical trials. New York: Raven Press, 1991.
- Gotzsche PC. Is there logic in the placebo? *Lancet* 1994; **344**: 925–26.
- Liberman RP. The elusive placebo reactor. *Neuropsychopharm* 1967; **5**: 557–66.
- Joyce CRB. Placebos and other comparative treatments. *Br J Clin Pharmacol* 1982; **13**: 313–18.
- Moerman DE. General medical effectiveness and human biology: placebo effects in the treatment of ulcer disease. *Med Anthropol* 1983; **14**: 3, 13–16.
- Beecher HK, Keats AS, Mosteller F, Lasagna L. The effectiveness of placebo “reactors” and “non-reactors”. *J Pharmacol Exp Ther* 1953; **109**: 393–400.
- Dinnerstein AJ, Lowenthal M, Blitz B. The interaction of drugs with placebos in the control of pain and anxiety. *Perspect Biol Med* 1996; **10**: 103–117.
- Wolf S. Effects of suggestion and conditioning on the action of chemical agents in human subjects—the pharmacology of placebos. *J Clin Invest* 1950; **29**: 100–09.
- Marks HM. The progress of experiment: science and therapeutic reform in the United State, 1900–1990. Cambridge: Cambridge University Press, 1997.
- Lilienfeld AM. *Ceteris paribus*: the evolution of the clinical trial. *Bull Hist Med* 1982; **56**: 1–18.
- Kaptchuk TJ. Intentional ignorance: a history of blind assessment in medicine. *Bull Hist Med* 1998; in press.
- Hill AB. Suspended judgment: memories of the British streptomycin trial in tuberculosis. The first randomized clinical trial. *Control Clin Trials* 1990; **11**: 77–79.
- Hill AB. The clinical trial. *N Eng J Med* 1952; **274**: 4 113–19.
- Batterman RC, Grossman AJ. *J Am Med Assoc* 1955; **159**: 17 1619–22.
- McDonald CJ, Mazzuca SA, McCabe GP. How much of the placebo “effect” is really statistical regression? *Stat Med* 1983; **2**: 417–27.
- Kienle GS. Der sogenannte placeboeffekt: illusion, fakten realität. Stuttgart: Schattauer, 1995.
- Kienle GS, Kiene H. Placebo effect and placebo concept: a critical methodological and conceptual analysis of reports on the magnitude of the placebo effect. *Altern Ther* 1996; **2**: 39–54.
- Tyler DB. The influence of a placebo, body position and medication of motion sickness. *Am J Phys* 1946; **146**: 458–66.
- Diehl HS. Medicinal treatment of the common cold. *J Am Med Assoc* 1993; **101**: 2042–49.
- Gold H, Kwit NT, Otto H. The xanthines (theobromine and aminophylline) in the treatment of cardiac pain. *J Am Med Assoc* 1937; **108** (26): 2173–79.
- Gay LN, Carliner PE. The prevention and treatment of motion sickness. *Bull Johns Hopkins Hospit* 1949; **84**: 470–87.
- Sullivan MD. Placebo controls and epistemic control in orthodox medicine. *J Med Philos* 1993; **18**: 213–31.
- Kemphorne O. Why randomize? *J Stat Plan Inf* 1977; **1**: 1–25.
- Evans FJ. The placebo response in pain reduction. *Adv Neurology* 1974; **4**: 289–96.
- Roberts AH, Kewman DG, Mercier L, Hovell M. The power of nonspecific effects in healing: implications for psychosocial and biological treatments. *Clin Psychol Rev* 1993; **12**: 375–91.
- Porter DR, Capell HA. The ‘natural’ history of active rheumatoid arthritis over 3–6 months: an analysis of patients enrolled into trials of potential disease-modifying anti-rheumatic drugs, and treated with placebo. *Br J Rheumatol* 1993; **32**: 463–66.
- McQuay H, Carroll D, Moore A. Variations in the placebo effect in randomized controlled trials of analgesics: all is blind as it seems. *Pain* 1995; **65**: 331–35.
- Kleijnen J, de Craen AJM, van Everdingen J, Krol L. Placebo effects in double-blind clinical trials: a review of interactions with medications. *Lancet* 1994; **344**: 1347–49.
- Ernst E, Resch KL. Concept of true and perceived placebo effects. *BMJ* 1995; **311**: 551–53.
- Kaptchuk TJ, Eisenberg DM. Legitimate evidence: randomized or otherwise. *Advances* 1997; **13**: 48–51.