IJE vol.33 no.2 © International Epidemiological Association 2004; all rights reserved. *International Journal of Epidemiology* 2004;**33**:247–251
DOI: 10.1093/ije/dyh047

Commentary: Unbiased divination, unbiased evidence, and the patulin clinical trial

Ted J Kaptchuk[1] and Catherine E Kerr[2]

Histories of the randomized clinical trial (RCT) are constantly unearthing dramatic precursors for concurrent controls, random assignment to comparison group, blind assessment, placebo controls, and statistical inference.[1–4] All these forerunners seem to indicate a constantly budding but, somehow, stumbling and aborted development toward the modern methodology. While most of the methods (at least in primitive forms) seem known in numerous early sources, it is not clear why medicine as a profession refused to commit itself to safeguards against bias until after World War II. The 60th anniversary of the under-appreciated and critical placebo controlled clinical trial of patulin for the common cold is a timely moment to offer some historical reflections on the origin and final acceptance of the RCT.

This commentary will hypothesize that the modern methods to reduce bias in clinical research—especially randomization, concurrent controls, and blinding—in at least rudimentary forms—were well known to educated physicians and, in fact, most literate European and Middle Easterners for thousands of years. Most of these assurances against bias or chicanery have their origin in religious divination rituals. From the 16th century onwards until the 1950s, many physicians sought to adopt these methods for revealing secular and unbiased medical truths. Despite valiant efforts by medical reformers, the bulk of the medical community ignored or rejected the introduction of unbiased comparison of clinical outcomes. Until 1950, most of the medical profession was not attracted to any form of clinical experimentation that challenged the personal judgement of physicians or seemingly threatened to treat patients as less than unique individuals. 'Science' was thought best confined to the laboratory where variability could be contained and physicians were best suited to delicately apply this knowledge within the context of 'the art of healing'.

The 1944 patulin trial and its more famous younger sibling, the 1948 streptomycin trial for tuberculosis, were both deeply embedded in the British biometry statistical tradition. These trials were not so important for their methodological innovations, as much as they came to represent potent exemplary models for an innovative team of statisticians and epidemiologists who were disseminating already known methodologies and convincing colleagues of the need and the possibility for a 'clinical science'. This re-construction of medicine's self-identity was as great a revolution as any previous transformation in the history of medicine. The patulin trial was a key component of this revolution as well as an important illustrative example of the very process itself.

Unbiased divination

In religious traditions, miracles and supernatural occurrences are the primary vehicle for divine revelation. However, waiting for miracles can be frustrating or too slow, and religions often develop methods to initiate divine communications on request. Attempting such connection on demand—rituals that expect direct answers from the deities—is to walk a treacherous slope. The monotheistic religions especially worried that divination techniques somehow represented magic that might compromise or restrict God's absolute omnipotence.[5,6] Nonetheless, most religions, even with the complex ambivalence (and sometimes absolute opposition) of the monotheistic traditions made room for divine contact on appeal. Because such theurgic manoeuvres involved human intentionality, divination was considered especially susceptible to prejudice or chicanery.[7] To prevent the distortions of human manipulation, divination often incorporated various 'safeguards' that insured against tampering by humans.

Thus, although both the Jewish and Christian Bibles often criticize divination as an abomination and sorcery,[8] (especially if performed by alien religions) (Leviticus 19:25, Deuteronomy 18:9–13) episodes in which chance-based mechanisms are used to elicit divine guidance are extremely common. For example, the Bible often assumes that a method of random selection, such as throwing lots, prevents human consciousness from distorting heavenly directives. Proverbs (16:33) states: 'Lots are cast into the lap; the decision depends on the Lord.'[9] Chance is out of human control. Countless examples can be found in the Hebrew Bible where lots allow accurate and uncorrupted communication of divine intention and demonstrations of truth. The priests in Jerusalem's temple wore a 'breast-piece of decision' which contained inscribed oracle discs (the Urim and Tummmim) that answered yes-no questions by a lottery mechanism. (Exodus 28:28–30, c.f. Numbers 27:19–21, Leviticus 8:8, I Samuel 14:37–42) Poorer people sometimes seemed to settle for rhabdomancy (divination with sticks). (Hosea 4:12) Lots determine the choice of the scapegoat for the Day of Atonement (Yom Kippur) sacrifice. (Leviticus 16:8) Purim, the holiday established in the Book of Esther, derives its name from the Hebrew word 'pur' meaning lots. (Esther 9:24) When the ship carrying a fleeing Jonah is threatened with stormy destruction, the sailors cast 'lots' to find out the true object of the Lord's anger.

[1] Osher Institute, and [2]Department of Social Medicine, Harvard Medical School, 401 Park Drive, Landmark Building, Boston, MA 02215, USA.

Correspondence: Ted J Kaptchuk, Osher Institute, Harvard Medical School, 401 Park Drive, Landmark Building, Boston, MA 02215, USA. E-mail: Ted_Kaptchuk@hms.harvard.edu

(Jonah 1:7) Later rabbis comment that the plural word 'lots' is used because a single lot draw might have to do with simple bad luck.[10] Similarly, the Christian New Testament reports lots being used to find a replacement for Judas after his betrayal. (Acts 1:23–26) And despite the especially vociferous condemnations of lots and divination from medieval Church scholastics,[11] no amount of disputation could hide from educated Christians St Augustine's use of rhapsodomancy (randomly opening a sacred text) during a critical moment of his spiritual awakening and conversion.[12] In his moment of uncertainty, Augustine only trusted the page he randomly selected from the Bible. For a long time, Europeans knew that chance mechanisms (best performed by someone with the correct religious credentials) could reveal 'truth' independent of human bias. (Obviously, the Hebrew and Christian use of chance has earlier historical precedence; Mesopotamian, Assyrian, and Egyptian priests all consulted lots [as did the diviners of many other cultures].)[13,14]

The Hebrew Bible also employs comparative trials (in some ways similar to modern trials) designed to reveal untainted truth. The most famous example is in Chapter 1 of the Book of Daniel. King Nebuchadnezzar of Babylon decides to train Daniel and three other exiled Judeans to become courtiers. Presumably to avoid breaking kosher laws, Daniel requests permission from the king's officer 'not to defile himself' with the 'daily rations … from the king's food and from the wine he drank.'[15] (Daniel 1:5) The officer, while kindly disposed to Daniel, feared that the King would notice that Daniel and friends would look 'out of sorts, unlike the other youths of your age—and … put [his] life in jeopardy'. (Daniel 1:10) Therefore, Daniel offers to do a prospective comparative test: his group would eat legumes and water and the rest of the trainees would eat the king's food. After a test of 10 days, Daniel 'looked better and healthier than all the youths' eating the regal faire. (Daniel 1:15) Eventually in the 'follow-up', keeping simple dietary rules seemed to make the Judeans ' 10 times better [in wisdom] than all the magicians and exorcists throughout the realm.' (Daniel 1:20) The Prophet Elijah also performs a comparative experiment, which included elaborate precautions against bias and fraud, with the priests of Baal to see whose worship could ignite a sacrificial altar. (I Kings 18:19–40)

Literate European readers would also have been aware of the use of randomization and even blinding from other traditions. For example, the Iliad has many incidents of lots being used to select opponents for battle, including a case where Hector shakes the lots and 'looks backward' so as to avoid appearing to manipulate the outcome.[16] More importantly, the centrality of oracles for Greco-Roman religions was well-known as were the particulars of the chance mechanism in use at Delphi or the Temple of Fortuna (where lots selection was sometimes augmented by using children to draw the lots to further prevent manipulation).[17,18]

Unbiased experimentation

As the modern era emerged—and secular knowledge assumed increasing importance—it is not surprising that physicians with challenges to established medical dogma appealed to well-known impartial methods to overcome human prejudice. Thus, from early on, starting with medieval Islamic sources, (Table 1) it is easy to find numerous (probably proposed) clinical trials, not unlike the comparisons of the Prophets Daniel and Elijah (usually proposing large numbers of patients for additional emphasis), strewn across well-known literature.

Slightly later examples of actually performed experiments designed to challenge established dogma are well-known. The earliest formal modern experiment we have found is Francis Bacon's (1627) elaborate investigation of the effects of steeping wheat seeds in various solutions in a search for the most effective means to hasten germination.[22] Bacon describes in detail his division of the seeds into separate groups, with each group steeped in water primed with either urine, one of several different types of animal dung, wine, or chalk. Other examples

Table 1 Early examples of medical discussions of unbiased comparisons

Date	Author	Proposed experiment
900–1100AD	**Al Rhazi and Ibn Sina,** Islamic physicians.[19]	Various passages in the writings of Islamic medical texts describe comparative tests of various interventions including bloodletting.
14th century	**Petrarch,** Renaissance classicist, poet and scholar.[3]	Describes a physician friend's proposal for a comparative test of 'a hundred or a thousand of men' with any identical disease half treated by physicians and half treated by 'Nature' alone.
1537	**Ambroise Paré,** French barber-surgeon.[20]	Describes an improvised (probably accidental) experiment in which about half of a large group of wounded soldiers receive cauterizing with hot oil and the other half receive an ointment of eggs, oils of roses, and turpentine.
1662	**Van Helmont,** Flemish iconoclastic and physician-chemist.[1]	Proposes an experiment comparing 200 or 500 people with fevers and pleurisy divided by 'lots' into a bloodletting/purging group and a group treated with his new chemical methods. The outcome would be number of funerals.
1752	**Bishop Berkeley,** prominent philosopher and natural scientist.[21]	Proposes an experiment of equal numbers of smallpox patients in hospital situations matched for diet, lodging, and time of year to compare the standard drugs and his unconventional use of tar water.
1784	**Franz Anton Mesmer,** Viennese physician who claimed to discover 'animal magnetisim'.[2]	Proposes an experiment where an equally large number of patients with any disease except venereal ones are treated with his methods and compared with those treated by regular physicians.

of medical experiments follow. Smallpox and later cowpox inoculation seems to be the impetus for numerous early prevention comparisons,[23–25] and, by 1760, Daniel Bernoulli had applied probability theory to inoculation outcomes.[26] Of course, the classic and most well-known early deliberately planned concurrent comparative trial seems to be James Lind's 1747 study of eight sailors with scurvy which evolves out of an emerging British understanding of comparative analysis.[27]

It is only in the late 18th century onwards, when elite medicine bifurcates and becomes enmeshed in a tug-of-war between mainstream and unconventional medicine, that the imperative to demonstrate medical outcomes uncorrupted by poor judgement, illusion, over-enthusiasm, imagination, and fraud becomes an urgent matter. Ongoing skirmishes between the two camps promoted the adoption and elaborate refinement of concurrent comparisons, random assignment, placebo controls, and statistical analysis.[2] Such placebo controlled trials (increasingly utilizing double blinding and larger numbers of subjects) were continually published in prominent medical journals for over a century prior to the patulin and streptomycin trials.[2] In these debates, fair methods were necessary for evaluating unconventional medicine but of no consequence for mainstream's self-evaluation.[2]

In France, a small group of physicians developed a concern for fair comparison for mainstream therapeutics that centred around the 'numerical method' of Pierre-Charles-Alexandre Louis, most famously developed in his experiments on bloodletting in 1828.[28] These experiments were quickly followed by Gavarret's ideas (based on Poisson's probability theories) which fostered the early creation of what we now call significance testing. Gavarret used a 'limits of oscillation' to determine whether the difference between two means was sufficient to rule out the effects of chance.[28]

Further methodological innovation can be seen emerging from mid-19th century experiments on the psycho-physics of sensory discrimination, telepathy, and suggestion/hypnosis. Psychologists developed a robust tradition of concurrent comparison, randomization, and placebo controls.[2,29,30] By 1860, Gustav Fechner, one of the pioneers of psycho-physics developed a method of significance testing based on 'average error.'[31] By the early 20th century, well before RA Fisher's work, psychologist-educators and even social reformers had worked out random methods for equalizing groups in comparative experiments and significance tests.[32,33]

Why were so many developments of unbiased research methods—the methodologies used in the unconventional medicine debates, the innovations of Lind, Louis, and Gavarret in mainstream medicine, and the efforts in psychology, education, and social reform—although well-known, almost unanimously ignored by mainstream medicine for so long? The short answer is that until after World War II, the science of medicine was understood to mean the discipline of the laboratory, either pathological anatomy or later physiology and bacteriology. Science was thought best kept segregated from the variability and uncontrolled environment of real patients that required 'art' and an appreciation of individual idiosyncrasy. 'Scientific medicine was a matter of applying, at the bedside, knowledge produced elsewhere.'[34] Clinical medicine of all the disciplines was least attracted to statistics and 'it was far from clear that the mean results for some large number of assorted trials in a hospital provided an appropriate basis for treating [the] individual.[35] Claude Bernard, the founder of modern physiology, repeatedly reinforced this point, as he noted in 1875, 'statistical methods. … can supply us only conjectures, probabilities; we can draw no certitude for the particular case.' (ref. 28, p. 27) Even after World War II many doctors inveighed against:

> patients [being] degraded from human beings into bricks in a column, dots in a field, or tadpoles in a pool; with the eventual elimination of the responsibility of the doctor to get the individual back to health.[36]

The patulin trial

During World War II, a committee of the Medical Research Council (MRC) decided to test the curative effects of a product extracted from an atypical penicillium, *Penicillium patulinum*, in the common cold. Philip D'Arcy Hart was organizing secretary and the statisticians included Major Greenwood. The trial was placebo controlled and an elaborate double-checking method was instituted to insure that the alternative assignment to patulin or placebo group was followed fastidiously. The concealment was also meticulous: to prevent guessing assignment category, eight different coded ampoules contained either patulin or placebo. The rationale for the design of the trial was elegant and persuasive. In all 1449 patients were treated; 1348 were analysed and there was no statistical difference in outcome between the 668 treated with patulin and 680 with placebo at 24 hours and 48 hours and the control was statistically better at one week.[37] (Obviously, later innovations such as intention-to-treat analysis and informed consent were missing.) While the patulin trial was methodologically impressive, the trial is often overlooked in historical discussions[38] with accolades usually reserved for the streptomycin for tuberculosis trial of 1948. The streptomycin trial actually used 'random sampling numbers' for treatment assignment (but abandoned patulin's placebo controls for pragmatic reasons).[39] Inaccurately, the streptomycin trial is often credited as being the first RCT in history. In fact, besides the unconventional trial and psychology experiments of the 19th century, by the 1930s, even within mainstream medicine, many controlled trials had already begun to use genuine random assignment methods (such as coin toss).[1,2]

Remembering the patulin trial, making it visible, is an important step in undoing the hagiographic history of the RCT. Both the patulin and streptomycin trials are not so much about new methodologies as they are emblematic milestones (especially taken together with the other MRC trials described below) in a genuine process of innovation, refinement, and, most importantly, diffusion initiated by the British biometry school. The patulin and streptomycin trials were rooted in the statistical tradition of such eminent statisticians as Francis Galton, Karl Pearson, Francis Ysidro Edgeworth, WS Gosset ('Student'), and RA Fisher. This school transformed statistics and probability into a tool for accurate measurement of variation, correlation, analysis of variance, and design of experiments. Their agenda was to transform phenomena with variability into processes accessible to hard scientific knowledge. Medicine was an important target in their revolutionary agenda. Well before patulin and streptomycin, Pearson and Greenwood (a student of Pearson)

discussed comparability of patients and methods for comparing medical intervention and controls.[40,41] 'Student' (Gosset) repeatedly discussed selection bias as a major threat in health experiments.[42] Followers of this school lead the debate in the MRC's Annual Report of 1928–1929 on whether there could be a genuine science that could cope with and tame the variability of real patients.[43] These statisticians and epidemiologists gradually adopted and refined methods already beginning to be used in mainstream experiments (and already well-established in psychology and unconventional medicine experiments). By 1934, MRC statisticians (including Greenwood) employed alternative selection in the trial of serum treatment of lobar pneumonia.[44] (And it should be noted that, at least since the time of Fibiger's diphtheria trial of 1898, alternative selection as a chance, allocation method was well known to mainstream medicine.[45]) The patulin trial, again under Greenwood's guidance, adopted placebo controls (already well-established elsewhere[2]) in addition to a rigorous and innovative method of alternative assignment. In 1946 Austin Bradford Hill (a student of Greenwood) convinced two MRC committees to use the method of 'random sampling numbers' for treatment assignment to further insure concealment.[1,46] Partly under D'Arcy Hart's directions, this innovative method of randomization was first used to create the comparison groups in the trial of pertussis vaccine (published in 1951[47]) and a few months later, to create the two arms in the streptomycin trial. Again the method of random assignment (without the refinement of using random numbers) already had a long history. (A fifth trial testing antihistamines for the common cold seems to be the first completed MRC double blind randomized trial and was both begun and reported in 1950.[48] This RCT was also under the direction of Hill and D'Arcy Hart and was actually the MRC's first published double-blind RCT.) These five trials taken together—serum treatment of pneumonia, patulin, streptomycin, pertussis vaccine, and antihistamines—are all linked sibling MRC experiments that mark the establishment, refinement, and dissemination of a model for a new clinical science.

All five main medical trials in the biometry tradition became beacons in a diffusion process of what many thought were 'new' methodologies. The intellectual powerhouse of the biometry school did not so much invent the RCT so much as root its rationale in sophisticated statistics and boldly disseminate it by enabling the medical profession's ability to conceptualize experimental testing in clinical situations. The work of the MRC group, especially after the publication of the streptomycin and antihistamine trials (guided by the intellectual acumen and persuasive abilities of Hill), quickly won allies with Beecher's and Lasagna's group at Harvard,[2] with Gold and Wolf's group at Cornell,[2] and with Cornfield's group at the National Institutes of Health (NIH) in the US.[49] Following Latour's portrait of 'science in action,' we can see that the MRC investigators, by enlisting and interesting a range of actors beyond the MRC, were able to broadcast the RCT as a necessary apparatus for the production of medical facts.[50] This manoeuvre significantly 'upset the balance between science and clinical experience' that had existed in medicine until that point. (ref. 43, p. 599)

Why was this group able to spread its influence so quickly and effectively when earlier researchers proposing similar methodologies in medicine saw their efforts founder? An entire paper would be necessary to begin to answer this question. Certainly, among the reasons would be: (1) the stream of new drugs requiring rational evaluation, (2) the new scientific legitimacy of probabilistic statistics after Maxwell's work on the molecular velocities of gas, (3) the desire of physicians to share in the iconographic power of science, (4) the institutional power of the MRC, and (5) the persuasive abilities of such men as Austin Bradford Hill. In this regard, one should also remember the legislative imprimatur given to this methodology after the thalidomide tragedy and the 1962 amendment to the American Pure Food and Drug Act (which finally gained the force of regulatory law in 1970).

A triumphalist history, which portrays methods for eliminating bias as unavailable until well into the 20th century when supposedly Austin Bradford Hill or the MRC (presumably under the influence of RA Fisher) magically introduce it into medicine, depicts the RCT as lacking a cultural, political, and social history. Remembering patulin reveals that history is complex, it involves incremental steps and that social networks can be critical. In fact, an examination of the RCT's genuine history (including aspects of that history described in this essay and elsewhere [e.g 1–4]) reveals an elaborate cultural, social, and scientific process. The patulin trial of 1944 reminds us that the struggle to produce untainted and uncorrupted evidence has a long and human (and even divine) history. Patulin was a key link in a long process that allowed the streptomycin trial of 1948 to become the celebrated seminal hagiographic moment for clinical research. In fact, the patulin trial was a crucial component of a family of trials created by MRC medical reformers seeking to transform medicine's central self-identity from art to science.[51] With this transformation accomplished, elite status shifted toward 'clinical scientist' and away from the 'experienced physician.' Few efforts have changed medicine so much.

References

[1] Chalmers I. Comparing like with like: some historical milestones in the evolution of methods to create unbiased comparison groups in therapeutic experiments. *Int J Epidemiol* 2001;**30:**1156–64.

[2] Kaptchuk TJ. Intentional ignorance: a history of blind assessment and placebo controls in medicine. *Bull Hist Med* 1998;**72:**389–433.

[3] Lilienfiled AM. Ceteris paribus: the evolution of the clinical trial. *Bull Hist Med* 1982;**56:**1–18.

[4] Dean ME. A homeopathic origin for placebo controls: 'an invaluable gift of God.' *Altern Ther Health Med* 2000;**6:**58–66.

[5] Thomas K. *Religion and the Decline of Magic*. New York: Charles Scribner's Sons, 1971.

[6] Neusner J. Science and magic, miracle and magic in formative Judaism: the system and the difference. In: Neusner J, Frerichs ES, Flesher PVM (eds). *Religion, Science and Magic*. Oxford: Oxford University Press, 1989.

[7] Evans-Prichard EE. *Witchcraft, Oracles and Magic among the Azande*. Oxford: Clarendon Press, 1976.

[8] Silverman WA, Chalmers I. Casting and drawing lots: a time honoured way of dealing with uncertainty and ensuring fairness. *BMJ* 2001;**323:**1467–68.

[9] New JPS Translation According to the Traditional Hebrew Text. *Tanakh: The Holy Scriptures*. Philadelphia: Jewish Publication Society, 1988, p. 1312. All further Hebrew Bible citations and quotes are from this translation.

[10] Rabinovitch NL. *Probability and Statistical Inference in Ancient and Medieval Jewish Literature*. Toronto: University of Toronto Press, 1973.

[11] Flint VIJ. *The Rise of Magic in Early Medieval Europe*. Princeton: Princeton University Press, 1991.

[12] Palkington JG (trans). *The Confessions of St. Augustine*. New York: Liveright, 1943.

[13] Bennett DJ. *Randomness*. Cambridge: Harvard University Press, 1998.

[14] David FN. *Games, Gods and Gambling: Origins and History of Probabilty and Statistical Ideas from the Earliest Times to the Newtonian Era*. New York: Hafner, 1962.

[15] New JPS Translation According to the Traditional Hebrew Text. *Tanakh: The Holy Scriptures*. Philadelphia: Jewish Publication Society, 1988, p. 1469.

[16] Lang A, Leaf W, Myers E. (trans) *The Complete Works of Homer. The Iliad and the Odyssey*. New York: Modern Library, 1938, p. 54.

[17] Burkert W. *Greek Religion*. Cambridge: Harvard University Press, 1985 [1977].

[18] Cicero MT. *De Divinatione*. Falconer WA (trans.). London: William Heinemann, 1928.

[19] Dallal A. Science, medicine and technology: the making of a scientific culture. In: Esposito JL (ed.). *The Oxford History of Islam*. Oxford: Oxford University Press, 1999.

[20] Keynes G. (ed.). *The Apologie and Treatise of Ambrose Paré*. New York: Dover 1968.

[21] Berkeley G. Farther thoughts on tar-water (1752). In: Sampson G (ed.). *The Works of George Berkeley, DD Bishop of Cloyne*. London: George Bell, 1898.

[22] Bacon F. *Sylva Sylvarum, or a Naturall History*. London: William Rawley, 1627, pp. 109–11.

[23] Miller G. *The Adoption of Inoculation for Smallpox in England and France*. Philadelphia: University of Pennsylvania Press, 1957.

[24] Fitz RH. Zabdiel Boylston, inoculation, and the epidemic of smallpox in Boston in 1721. *Bull Johns Hopkins Hospital* 1911;**22:**315–27.

[25] Boylston AW. Clinical investigation of smallpox in 1767. *N Engl J Med* 2002;**346:**1326–28.

[26] Daston L. *Classical Probability in the Enlightenment*. Princeton: Princeton University Press, 1988.

[27] Thöhler U. *'To Improve the Evidence of Medicine.' The 18th Century British Origins of a Critical Approach*. Edinburgh: Royal College of Physicians of Edinburgh, 2000.

[28] Mathews JR. *Quantification and the Quest for Medical Certainty*. Princeton: Princeton University Press, 1995.

[29] Dehue T. Deception, efficiency, and random groups. Psychology and the gradual origination of the random group design. *Isis* 1997;**88:**653–73.

[30] Hacking I. Telepathy: origins of randomization in experimental design. *Isis* 1988;**79:**427–51.

[31] Stigler SM. *The History of Statistics: The Measurement of Uncertainty before 1900*. Cambridge: Harvard University Press, 1986.

[32] Dehue T. Establishing the experimenting society: the historical origin of social experimentation according to the randomized controlled design. *Am J Psych* 2001;**114:**283–302.

[33] Oakly A. Experimentation and social interventions: a forgotten but important history. *BMJ* 1998;**317:**1239–42.

[34] Marks HM. *The Progress of Experiment. Science and Therapeutic Reform in the United States, 1900–1990.* Cambridge: Cambridge University Press, 1997, p. 1.

[35] Porter TM. *The Rise of Statistical Thinking 1820–1900*. Princeton: Princeton University Press, 1986, p. 157.

[36] Waugh WG. A blast of the trumpet against the monstrous regiment of mathematics. [Letter] *BMJ* 1951;**i:**1088.

[37] Medical Research Council Patulin Clinical Trials Committee. Clinical trial of patulin in the common cold. *Lancet* 1944;**ii:**373–75.

[38] D'Arcy Hart P. A change in scientific approach: from alternation to randomized allocation in clinical trials in the 1940s. *BMJ* 1999;**319:**572–73.

[39] Medical Research Council. Streptomycin treatment of pulmonary tuberculosis. *BMJ* 1948;**ii:**769–82.

[40] Pearson K. Report on certain enteric fever inoculation statistics. *BMJ* 1904;**ii:**1243–46.

[41] Greenwood M, Yule GU. The statistics of anti-tyhphoid and anti-cholera inoculations and the interpretation of such statistics in general. *Proc R Soc Med* 1915;**8:**113–20.

[42] 'Student' [Gosset WS] The Lanarkshire milk experiment. *Biometrika* 1931;**23:**398–406.

[43] Armstrong D. Clinical sense and clinical science. *Soc Sci Med* 1977;**11:**599–601.

[44] Medical Research Council. The serum treatment of lobar pneumonia. *Lancet* 1934;**i:**290–95.

[45] Hróbjartsson A, Gøtzsche, Gluud C. The controlled clinical trial turns 100 years: Fibiger's trial of serum treatment of diphtheria. *BMJ* 1998;**317:**1243–45.

[46] Doll R. Controls trials: the 1948 watershed. *BMJ* 1998;**317:**1217–20.

[47] Medical Research Council. The prevention of whooping-cough by vaccination. *BMJ* 1951;**i:**1463–71.

[48] Medical Research Council. Clinical trials of antihistaminic drugs in the prevention and treatment of the common cold. *BMJ* 1950; **ii:**425–30.

[49] Ederer F. Jerome Cornfield's contribution to the conduct of clinical trials. *Biometrics* 1982;**39:**25–32.

[50] Latour B. *Science in Action*. Cambridge: Harvard University Press, 1987.

[51] Magnello E. The introduction of mathematical statistics into medical research: the roles of Karl Pearson, Major Greenwood and Austin Bradford Hill. In: Magnello E, Hardy A (eds). *The Road to Medical Statistics*. Rodopi: Amsterdam, 2002.