The Emergence of the Randomized, Controlled Trial

The birth of the randomized, controlled trial (RCT) is typically dated to a 1948 evaluation by the British Medical Research Council (MRC) of streptomycin for the treatment of tuberculosis. But controlled clinical trials and discussions of their designs were increasingly being published in medical journals for at least half a century before the MRC's report, which was part of a much longer history of efforts to empirically assess experimental therapies. An exploration of this deeper history offers insights into the intellectual and social forces shaping both the advent of and resistance to the controlled clinical trial as a medical research standard and mechanism for taming the therapeutic marketplace.

Trials involving experimental and control groups seem as old as the historical record itself, appearing in the Hebrew Bible and in various societies around the world, albeit sporadically, for centuries.[1] As Enlightenment reasoning filtered into medicine, controlled trials emerged with growing frequency. In 1753, Scottish surgeon James Lind published a controlled trial demonstrating that a diet including citrus fruit was effective against scurvy in sailors at sea, thereby providing a touchstone for subsequent generations of researchers who gradually embraced comparative trial methods.

Loosely controlled trials increasingly appeared in the 18th and 19th centuries, often conducted by skeptics to test the utility of unorthodox remedies ranging from mesmerism to homeopathy.[2] Major shifts in the social and scientific structure of medicine in the late 19th and early 20th centuries created new opportunities and demands for more rigorous clinical research methods. Hospitals expanded, new biologic and vaccine industries emerged to confront recently identified germs, chemists developed novel therapeutic compounds, and an unregulated subeconomy of fraudulent replicas of new agents flourished. All these factors motivated clinical investigators to pursue more sophisticated approaches for evaluating experimental therapies.[3,4]

By the late 19th century, researchers were conducting "alternate-allocation" trials, the most recent methodologic ancestor of RCTs. Conventionally dated to Johannes Fibiger's 1898 study of diphtheria antitoxin in 484 patients in Copenhagen, alternate allocation entailed treating every other patient (or, in Fibiger's case, patients seen every other day) with a particular experimental remedy, withholding it from the others, and then comparing outcomes. But Fibiger's was only the most famous use of a technique that increasingly appeared in the medical literature from the 1890s onward, one that could (though only occasionally did) involve patient or researcher blinding, use of placebos for control groups, and statistical analysis of results.[1]

The pages of the *Journal* reflect this transition in research methods and the gradual but limited adoption of alternate-allocation trials to verify the purported value of new therapies. As early as 1899, a Dr. Williams described applying a glycerin–hydrogen peroxide solution "to the skin of

every alternate patient" to treat desquamation owing to scarlet fever, finding shortened periods of desquamation among treated patients (1899; see box for historical *Journal* articles). Over the ensuing half-century, medical journals published numerous primary reports of alternate-allocation studies (most of them of infectious disease treatments), along with discussions of such studies and appeals for using this method to resolve disputes (see Selected Alternate-Allocation Studies and Discussions Published in the *Journal* before 1948).1

Admittedly, in 1931, James Burns Amberson and colleagues published a study in which a coin flip randomly determined which of two seemingly equally divided groups of patients would receive sanocrysin for the treatment of tuberculosis. But this study was an outlier, with alternate-allocation studies serving as the dominant model of therapeutic controlled trials in the first half of the 20th century. The number of alternate-allocation studies, however, was itself dwarfed by the number of articles promoting therapies on the basis of other forms of evidence, from laboratory and physiological justifications to case reports. Many producers of new treatments lacked economic, regulatory, or social incentives to rigorously evaluate their products in controlled trials, and many researchers simply continued relying on standard methods that were widely accepted by scientists and society.

Some researchers resisted controlled trials because they believed that participants should not be denied promising treatments by being assigned to control groups. As an editorialist lamented in 1935 regarding a trial evaluating convalescent serum for the treatment of poliomyelitis, "Parents do not have to be persuaded or urged to volunteer their children for the trial of new biologic agents — they demand them. . . . [But] means for careful appraisal were easy to devise; impossible to carry out" (1935a). That trial's protocol called for administering the serum only in alternate cases, but researchers hesitated: "The main difficulty encountered was the inability of our special investigators to withhold this promising agent from any stricken child. . . . Our sentiment overruled our reason."

Alternate allocation also stimulated debates about methodology. Proponents argued that alternate allocation was superior to conventional case-study methods or expert testimonials for estimating an intervention's therapeutic value. As a 1936 editorial on the use of novel sulfa drugs cautioned, "the only way to evaluate properly a therapeutic agent of this sort is by the 'alternate case method' . . . even though statements of reputable physicians are impressive" (1936c). Critics, however, identified deficiencies in alternate allocation, as illustrated by a series of pneumonia trials. In 1924, Boston City Hospital's Edwin Locke alternately assigned pneumonia patients to receive or not receive antipneumococcal antiserum and found no difference in mortality between treated and untreated patients (1924). When Maxwell Finland repeated the study a few years later, he found a benefit for the serum but admitted that "the data reveal the possibility that some choice may have been unconsciously exercised in selecting cases for treatment" (1930).

Such concerns about selection bias stemming from the ease of cheating the process of strict allocation remained the Achilles' heel of alternate allocation and recurred in debates over a trial of diathermy for pneumonia in 1935 (1935b, 1935c, 1936b) and another trial comparing treatment with serum alone versus serum plus sulfa drugs for pneumonia in 1941. Again, Finland noted the shortcomings of the alternate-allocation schemes, pointing charitably to prevailing ethical concerns in surmising that "some unconscious selection on the part of the authors played

an important role in the inclusion of the poorest subjects among the serum recipients" (1941a, 1941b).

Such methodologic concerns resonated with British epidemiologist–statistician Austin Bradford Hill, who had grown familiar with alternate allocation's limitations while evaluating a series of MRC trials of antipneumococcal antiserum in the early 1930s. By the time he devised the MRC's evaluation of streptomycin for tuberculosis in the 1940s, Hill was sufficiently concerned about researchers' capacity to figure out (and hence cheat) allocation schemes that, in an attempt at frustrating such efforts, he replaced alternate allocation with strict concealed randomization of patients to treatment or control groups.[5] The blinding of researchers to patients' assignments, if at all possible, soon accompanied concealed random allocation in the emerging definition of the ideal study, in which bias was to be eliminated.[2]

Supported by MRC funding in the 1940s and 1950s, Hill and his colleagues impressed the research community with a series of groundbreaking RCTs. British investigators were soon followed by U.S. and other researchers who embraced RCTs as urgently needed tools for separating the wheat from the chaff emanating from an ever-diversifying pharmaceutical industry. When the U.S. Congress passed the 1962 Kefauver–Harris Amendments to the Food, Drug, and Cosmetic Act, the RCT had become an obvious methodology by which the Food and Drug Administration (FDA) could require pharmaceutical manufacturers to demonstrate therapeutic safety and efficacy before drug approval. By 1970, the FDA required that drug producers submit RCT results with new drug applications, heralding a new era with the RCT at the center of national, and expanding international, drug regulatory infrastructures.[4]

RCTs thus represent the most recent outgrowth of a long history of attempts to adjudicate therapeutic efficacy. Their immediate ancestor, alternate-allocation trials, emerged as part of a trend toward empiricism and systematization in medicine and in response to the need for more rigorous assessment of a rapidly expanding array of experimental treatments. Alternate allocation represented a significant advancement for addressing clinical research bias — but one that had limitations as long as it allowed foreknowledge of treatment allocation. Concealed random allocation emerged as the solution to these limitations, and RCTs were soon supported by crucial public funding and scientific regulatory infrastructures. Like alternate-allocation trials, though, RCTs expanded not only because they offered more sophisticated methods but also because they served a critical social function: screening experimental therapies before they were broadly distributed and clarifying the actual effects of medical innovations.

Selected Alternate-Allocation Studies and Discussions Published in the *Journal* before 1948

1899. Boston Society for Medical Improvement. 141:347–9.

1913. Park WH. Antitoxin administration. 168:73–7.

1924. Locke EA. The serologic treatment of lobar pneumonia. 190:196–203.

1930. Finland M. The serum treatment of lobar pneumonia. 202:1244–7.

1931a. Davis WE. The incidence of untoward symptoms following the intravenous injection of sodium tetraiodophenolphthalein in cholecystography. 205:534–6.

1931b. The clinical meeting of the Massachusetts General Hospital staff. 205:1319–20.

1935a. "Polio" vaccines. 213:687–8.

1935b. Diathermy in lobar pneumonia: preliminary report. 213:796–8.

1935c. King DS. Correspondence: diathermy in lobar pneumonia. 213:1324–5.

1936a. Curphey TJ, Solomon S. The therapeutic value of calcium salts in serum sickness. 214:150–3.

1936b. Wetherbee W. Correspondence: a discussion of Dr. Donald S. King's criticism. 214:174–5.

1936c. Prontylin and prontosil. 215:1311.

1939. Lanman TH, Heyl HL. Empyema in children. 221:1003–9.

1940. Cutts M, Burgess AM, Chafee FH. The treatment of lobar pneumonia with sulfathiazole and sulfapyridine. 223:762–4.

1941a. Finland M. Controlling clinical therapeutic experiments with specific serums: with particular reference to antipneumococcus serums. 225:495–506.

1941b. Combined serotherapy and chemotherapy in pneumonia. 225:514–5.

1947. Artificial sunlight treatment in industry. 236:374–5.