

The controlled clinical trial turns 100 years: Fibiger's trial of serum treatment of diphtheria

Asbjørn Hróbjartsson, Peter C Gøtzsche, Christian Gluud

The British Medical Research Council's trial of streptomycin for pulmonary tuberculosis, published in 1948,¹ has been proposed as the first randomised trial in which random numbers were used and allocation of patients was effectively concealed. Before 1948 several randomised trials had been reported,² but the method of randomisation was either not stated³ or was open to selection bias—for example, randomisation with use of a deck of cards.⁴

The earliest of these trials was published in 1898.⁵ It investigated the effect of serum treatment on diphtheria and was conducted by the Danish Nobel laureate, Johannes Fibiger. It was the first clinical trial in which random allocation was used and emphasised as a pivotal methodological principle. This pioneering improvement in methodology, combined with a large number of patients and rigorous planning, conduct, and reporting, makes the trial a milestone in the history of clinical trials.

Fibiger's trial was published in Danish and its method of randomisation has often been quoted incorrectly. We have translated central passages into English (available on the *BMJ* website at www.bmj.com) and discussed its methodological merit.

Fibiger and his trial

Johannes A G Fibiger (1867-1928) was born in Silkeborg, Denmark (figure). After receiving his medical degree in 1890 from the University of Copenhagen he visited Robert Koch and Emil von Behring in Germany. In 1895 Fibiger was awarded a doctoral degree for a thesis on diphtheria from the University of Copenhagen, where he became professor of pathology in 1900. He received the Nobel Prize in 1927, six weeks before he died, for research indicating that a nematode, *Spiroptera carcinoma*, caused gastric cancer in rats. The hypothesis was later rejected but, none the less, had an important role in the development of experimental research on cancer.^{6,7}

Fibiger conducted his diphtheria trial in 1896-7, when he worked as a junior physician at Blegdamshospitalet in Copenhagen.⁵ In the 1890s the evidence for an effect of serum treatment on diphtheria was contradictory. Tests on animals had been promising, and results from trials with concomitant controls performed by Roux in Paris had been convincing.⁸ Fibiger's superior at Blegdamshospitalet, Professor Sørensen, however, had not seen any effect of serum treatment in a series of trials.⁹ Furthermore, the side effect of serum treatment, serum sickness, was well known and strongly emphasised by the sceptics: "doctors refused serum treatment, although they themselves were ill with diphtheria."⁵

In his introduction, Fibiger makes it clear why he was not convinced by earlier investigations. Roux had compared patients treated with serum in one hospital with untreated patients in another hospital. The serum treatment was introduced at the same time as improve-

Summary points

A large randomised clinical trial was performed as early as 1898

Random allocation was emphasised as a central methodological principle

Patients were allocated to serum or no serum according to day of admittance, which created two comparable groups

The planning, conduct, and reporting of the trial was of high quality

The efficacy of serum treatment on diphtheria was shown

The trial was the first properly conducted controlled clinical trial

Nordic Cochrane Centre, H:S Rigshospitalet, Tagensvej 18 B, DK-2200 Copenhagen N, Denmark

Asbjørn Hróbjartsson, *PhD student*
Peter C Gøtzsche, *director*

Copenhagen Trial Unit, Centre for Clinical Intervention Research, Institute of Preventive Medicine, Copenhagen University Hospital, H:S Kommunehospitalet, DK-1399 Copenhagen K
Christian Gluud, *chief physician*

Correspondence to: Dr Gluud
cgluud@ipm.hosp.dk

BMJ 1998;317:1243-5

ments in isolation routines and hygiene, which meant that "the evidential weight of the experiments was lost." Sørensen concluded on the basis of his own work that "the trials were too few and the important number, the number of deaths, so small, that the material could hardly be used as a statistical proof; furthermore, the patients were allocated to serum treatment or the usual treatment by subjective judgment."⁵

The then 28 year old Fibiger convinced his professor that a new trial was necessary: "It was evident that a new series of experiments had to be planned and planned in such a way that the result would be absolutely conclusive." The objective of the trial was to investigate the effects of serum treatment on the mortality and morbidity of patients suffering from diphtheria. From 13 May 1896 to 13 May 1897 patients admitted to Blegdamshospitalet with diphtheria were treated with subcutaneous injections of diphtheria serum twice daily until symptoms improved in addition to standard treatment or with standard treatment alone.

Treatment allocation depended on the day of admittance. New patients admitted on alternate days received either standard treatment or standard treatment plus serum treatment. Patients were secondarily excluded if diphtheria bacteria were not identified. All patients had their throats swabbed with solutions of silver nitrate or tar oil. The main outcome was mortality; other outcomes were croup, dislodgement of membranes, temperature, albuminuria, and paralysis. The prevalence and intensity of serum sickness was also recorded.

Eight out of 239 patients in the serum treated group and 30 out of 245 in the control group died. No formal statistical analysis was performed but "no objection can



The authors' translation of central parts of Fibiger's original paper is available on the website

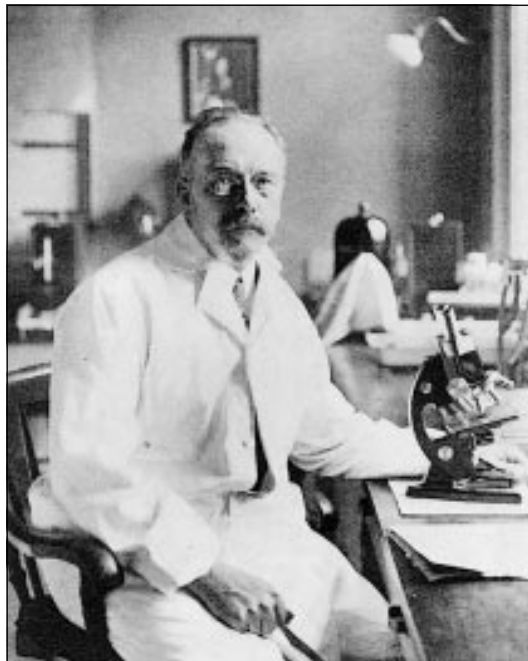
be raised against the statistical significance of the numbers," which were deemed correct by an inspector of the sick benefit association (in 1900 Pearson invented the χ^2 test,¹⁰ which would have shown $P=0.0003$). The rate of serum sickness was high at 60%.

Randomisation as the central methodological principle

Fibiger emphasised four methodological features of the trial: "Even with minimal knowledge of diphtheria epidemics, one will realise that it is necessary to have, firstly, large numbers and, secondly, a long study period." Thirdly he stated that "To compensate for the large seasonal variation in mortality, the study should last at least one year." His fourth contribution was a concise discussion of the importance of random allocation and bias. "Truly, the control cases in the earlier studies were selected to be as similar as possible to the ones treated with serum, but to eliminate completely the play of chance and the influence of subjective judgment, one had to use a different procedure. The only method which could be used rationally was to treat some patients with serum and every other patient in the usual way."

Thus, Fibiger took into account the two major threats to the validity of clinical trials: random error and bias. Fibiger had no illusions about clinical research methods of the time. He knew that investigations with historical control groups were generally untrustworthy. He also agreed with Sørensen that not only was a high number of patients necessary for a valid conclusion but, more importantly, a sufficiently high number of primary outcome events had to occur.

Fibiger had an impressive and clear view of bias associated with "subjective judgment" in patient allocation: "That this played not a trivial role can hardly be doubted... In many cases a trustworthy verdict can only be reached when a large number of randomly



Johannes Andreas Grib Fibiger (1867-1928)

Comparability of the two groups of patients with diphtheria in Fibiger's trial created by allocation to treatment according to day of admission

| Detail | Serum | No serum |
|---------------------------------|-------|----------|
| No of patients | 239 | 245 |
| No of adults (>15 years) | 56 | 58 |
| No affected by severe membranes | 73 | 88 |
| No with croup | 35 | 43 |

selected patients are treated with the new remedy and, at the same time, an equally large number of randomly selected patients are treated as usual." Therefore, "I suggested to Professor Sørensen to treat all patients admitted on the one day with serum, but none of those who were admitted the following day." Today, we know that this method is not immune to bias—for example, the practitioners could have admitted the milder cases preferentially on days with serum treatment. Such bias probably did not occur, however, as group sizes and baseline values for age and severity of disease were comparable (table).

Quality of the trial

By modern standards the general quality of Fibiger's trial is high. It was rigorously planned and conducted and was reported with clarity and detail. The outcomes of the study were clinically relevant, inclusion and exclusion criteria were clearly stated, and the flow of patients was reported. Of the 1004 patients admitted to Blegdamshospitalet with presumed diphtheria, 493 were "excluded from statistics, which of course comprise only diphtheria patients in whom diphtheria bacilli were identified. However, not all of those [with diphtheria bacilli] are included, some are excluded, partly those in whom the illness on admission was complicated by measles, scarlet fever or erysipelas, partly patients who on admission were moribund or in agony and died within 24 hours after admission (17 patients). An additional three patients, who were admitted with diphtheria bacilli in the throat but did not show any objective signs of diphtheria, are excluded. Furthermore, we have to withdraw from our statistics a few more patients, partly those who refused serum treatment (three mild cases), although they should have been injected according to the rota, partly two severe cases who were injected with serum according to the wish of the parents and the referring doctor, though they arrived on 'non-serum days,' and finally, near the end of the trial, a couple of seriously affected patients who were also treated with serum against the rota, as it was considered proven at this stage that the serum possessed a favourable effect." Such accuracy in reporting is rare today.

Fibiger was aware of the uncertainty of clinical judgment and explicitly defined some of the clinical outcomes. He also tried to minimise variation between observers by using "concordant observations by the consultant and myself."

Cointervention was a potential problem as the trial was not blinded, and severe cases in the control group were treated with mercury oxycyanide and mercury cyanide. This was probably without importance, however, as contemporary pharmacopoeias show that the concentrations of cyanide solutions were mostly weak (E Dauv-Pedersen, personal communication).

The wider context

Fibiger's trial had an immediate practical consequence. The demand for the new treatment increased to a point such that a whole new institute, the Serum Institute, had to be built. Unfortunately, Fibiger's methodological innovation had surprisingly little impact. The importance of random allocation was first fully recognised after the contributions of Fisher in 1925¹¹ and Bradford Hill in 1948.¹

Fibiger's conclusion was later criticised. Suspicion of inefficiency of serum treatment was raised during an epidemic in the 1920s, but the critique was based on uncontrolled observations. In 1951 Rasch re-analysed Fibiger's data and concluded that the effect was not significant.¹² Rasch's analysis was influential at the time but has recently been rejected.¹³ Fibiger emphasised the overall result, but he also divided his material into subgroups according to age and to the presence or absence of croup. Rasch based his critique on a subgroup of children below 6 years of age without croup. There is no a priori reason, however, why the outcome in the various subgroups should be different, and, in fact, the odds ratio for mortality for all cases is 0.34 versus 0.38 for the subgroup of children below 6 years of age.

In historical accounts Fibiger's trial is surprisingly often not quoted or misquoted. Patient allocation is usually described as alternation, treating every second patient with serum^{14 15} rather than treating patients according to day of admittance. It has also been argued that the trial was less valuable than it might have been as the diphtheria epidemic in 1896-7 had a low mortality.¹⁴ The study had sufficient power, however, and the number needed to treat to save one life was only 11, which is impressive. The most unique feature of the trial—the thorough discussion of avoidance of selection bias by random allocation—is rarely mentioned in the literature.

It is an interesting question as to who inspired Fibiger to develop the controlled clinical trial, which, by modern standards, would be called a quasirandomised trial. In Denmark Sørensen had conducted a clinical trial some years earlier in which the allocation in the final stage of the trial was based on alternation, but the principle was abandoned because "the method was difficult to carry out, and a subjective influence could not be excluded."¹⁶ Alternation was mentioned by Wanscher in 1877¹⁶ and emphasised by Heiberg in 1897,¹⁷ but none of them discussed randomisation as such or conducted a trial on the basis of random allocation.

Internationally, randomisation was mentioned by van Helmont in 1662.¹⁸ He suggested the casting of lots to decide which patients with fever should be assigned to blood letting and which should not. Randomisation was not used in clinical experiments, however, before 1898.¹⁹ In 1884 Peirce used a deck of cards to determine the order of weight changes in an investigation of the threshold for perception for small differences in weights.²⁰ Peirce's innovation was not appreciated, and he is mostly remembered as one of the founders of the philosophical movement called pragmatism. Some months later, the French Nobel laureate Richet published a series of trials that investigated whether thoughts could be transferred from one subject to another.²¹ After a person had drawn a card from a deck



Serum treatment for diphtheria, 1894

and concentrated on it another person should guess the suit of the card. Peirce and Richet did not reflect on the concept of randomisation in their articles.

Experiments were routine in laboratories. Fibiger, who was trained in the rigorous control of the microbiological experiment, paired its logic with the tradition of observational studies and created a new and important tool: the controlled clinical trial.

Contributors: CG did most of the literature search, AH wrote the first draft. All authors contributed to the final manuscript, reviewed central papers independently, and are guarantors.

Funding: No specific funding.

Competing interests: None declared.

- 1 Medical Research Council. Streptomycin treatment of pulmonary tuberculosis. *BMJ* 1948;ii:769-82.
- 2 Clarke M. Early controlled trials. *BMJ* 1996;312:1298.
- 3 Diehl HS, Baker AB, Cowen DW. Cold vaccines. An evaluation based on a controlled study. *JAMA* 1938;111:1168-73.
- 4 Tylor DB. The influence of a placebo, body position and medication on motion sickness. *Am J Physiol* 1946;146:458-66.
- 5 Fibiger J. Om Serumbehandling af Difteri. *Hospitalstidende* 1898;6:309-25, 337-50.
- 6 Stolley PD, Lasky T, Johannes Fibiger and his Nobel prize for the hypothesis that a worm causes stomach cancer. *Ann Intern Med* 1992;116:765-9.
- 7 Clemmesen J, Johannes Fibiger. Gongylonema and vitamin A in carcinogenesis. *Acta Path Microbiol Scand* 1978;suppl 270:1-13.
- 8 Roux ME, Martin ML, Chaillou MA. Trois cents cas de diphtérie traités par le sérum antidiphtérique. *Ann Inst Pasteur* 1894;8:640-62.
- 9 Sørensen S. Forsøg med Serumterapi ved Difteritis. *Hospitalstidende* 1896;4:621-8.
- 10 Pearson K. On the criterion that a given system of deviations from the probable in the case of a correlated system of variables is such that it can be reasonably supposed to have arisen from random sampling. *Phil Mag Ser* 1900;50:157-75.
- 11 Fisher RA. *Statistical methods for research workers*. Edinburgh: Oliver and Boyd, 1925.
- 12 Rasch G. Om kausal vurdering af terapi. *Ugeskr Læger* 1951;113:1625-30.
- 13 Gotzsche PC. Fibigers forsøg om serumbehandling af difteri. *Bibl Læger* 1997;189:143-6.
- 14 Bull JP. The historical development of clinical therapeutical trials. *J Chron Dis* 1959;10:218-48.
- 15 Lilienfeld AM. The evolution of the clinical trial. *Bull Hist Med* 1982;56:1-18.
- 16 Wanscher O. *Om diphtheritis og croup særlig med hensyn til tracheostomi*. København: Jacob Lund, 1877.
- 17 Heiberg P. Studier over den statistiske undersøgelsesmetode som hjælpemiddel ved terapeutiske undersøgelser. *Bibl Læger* 1897;7:1-40.
- 18 Van Helmont JA. *Oriatrike, or physick refined: the common errors therein refuted and the whole are reformed and rectified*. London: Lodowick-Lloyd, 1662:526. (Translated by Chandler J)
- 19 Hacking I. Telepathy: origins of randomisation in experimental design. *ISIS* 1988;79:427-51.
- 20 Peirce CS, Jastrow J. On small differences of sensation. *Memoirs Natl Acad Sci* 1884;3:73-83.
- 21 Richet C. La suggestion mentale et le calcul des probabilités. *Revue Philosophique de la France et de l'Étranger* 1884;18:609-74. (Accepted 5 October 1998)