

Tröhler U (2010). Adolf Bingel's blinded, controlled comparison of different anti-diphtheritic sera in 1918.

© Ulrich Tröhler, Institute of Social and Preventive Medicine (ISPM), University of Bern, Finkenhubelweg 11, CH-3012 Bern, Switzerland, Email: utroehler@ispm.unibe.ch

With the uptake of serum treatment for diphtheria from the end of the 19th century onwards, deaths from the disease fell dramatically (Russell 1943), albeit, less dramatically in countries, such as the UK, in which serum treatment had not been adopted wholeheartedly (Weindling 1992 p 73). The development of serum therapy of humans – in a race between the Berlin Institute for Infectious Diseases and Emile Roux and his colleagues at the Pasteur Institute in Paris - led to the award of the first ever Nobel Prize in physiology or medicine, in 1901, to Emil Behring. This illustrated that there was wide acceptance that anti-diphtheritic serum was effective.

Cite as:

Tröhler U (2010). Adolf Bingel's blinded, controlled comparison of different anti-diphtheritic sera in 1918. JLL Bulletin: Commentaries on the history of treatment evaluation (www.jameslindlibrary.org).

The accepted account of the serum's mechanism of action was that anti-toxins produced by artificially infected animals neutralised toxins released by diphtheria bacteria in infected humans (measured using a method developed by Ehrlich to measure anti-toxin levels in serum). Acceptance of the theory had been reflected in the establishment of 'healing-serum' factories all over the western world, using horses (because they are large animals) as 'anti-toxin producers'.

This orthodox explanation of the mechanism of action was challenged by the results of further clinical trials done by Adolf Bingel, a senior physician at the district general hospital of Brunswick, Germany (Bingel 1918). During a severe epidemic of diphtheria during the winter of 1910-1911, Bingel had become sceptical of the orthodox view that these specific 'healing-sera' were responsible for serum's anti-diphtheritic effects (Bingel 1918, p 284-285): variable responses to the anti-toxin serum had prompted him to wonder whether its beneficial effect was due solely to the anti-toxin. Might it not be caused by a non-specific action resulting simply from administering to patients a serum from another species? Bingel was aware of "the enormous influences of [foreign] protein from strange [non-human] species", as manifested in serum-disease, and its marked effects on haematological indices. He was at pains to emphasise that he had not the slightest intention of casting doubts on the results of animal research on immunity, but noted that the variable clinical picture of human diphtheria, "with its numerous and diverse complications", was completely different from the "experimentally induced infection or intoxication of an animal," so "whether a drug influences a human disease can only be decided in man." This reasoning of species-specific pathophysiology and therapy were the basis for Bingel's decision to undertake a study comparing (Behring's) 'anti-toxin serum' with 'normal serum', that is, serum derived from horses that had not been infected with diphtheria: "If no differences are found, the anti-toxin cannot be the effective agent" (Bingel 1918, p 284-285).

Given the widely acknowledged effectiveness of Behring's anti-toxin serum, Bingel proceeded cautiously, using alternation to create comparable groups of patients:

After I had treated some adult diphtheria patients with ordinary horse serum in 1911, I began in 1912 to treat alternate adult patients with antitoxin serum and with ordinary serum, exactly in the temporal sequence in which they were admitted to the ward. The children all received antitoxin serum. In the second half of the year 1912 and in the first half of 1913, I gradually lowered the age of those to be treated with ordinary horse serum, and from 1 July 1913 [till 31st December 1916 when he stopped the study], every second case was treated with ordinary horse serum, whether child or adult, regardless of the severity of the illness or the presence of complications.

Bingel noted that:

...it is absolutely inadmissible to compare the results for different time periods, for example to give antitoxin serum during one year, and then to give only ordinary horse serum during a second year, and then to

compare the results. That would lead to seriously wrong conclusions, for in no infectious disease is the nature of the epidemic so changeable as in diphtheria. Mostly we see light epidemics, but quite serious ones still occur. I remind [the reader] of the heavy epidemics in Berlin and Hamburg of the year 1910, and the one in Leipzig of 1914, which recall the bad times of the period before serum. (Bingel 1918, p 285-286).

In addition to using alternation to address this problem, Bingel also took steps to reduce observer biases. He noted that it was “extraordinary difficult...to evaluate the influences of therapy on disease unless they are obvious, as for example, the success of a surgical operation or cure of syphilis with mercury or Salvarsan. The therapeutic optimist very easily sees improvement, and the sceptic sees nothing.” In order to reduce these problems, Bingel concealed the identity of the two sera from his assistants and nurses, using the ‘cover names’ of ‘old serum’ (for the antitoxin serum) and ‘new serum’ or ‘red serum’ (for normal serum).

To make the trial as objective as possible, I have not relied on my own judgment alone, but have sought the views of the assistant physicians of the diphtheria ward, without informing them about the nature of the serum under test (namely the ordinary horse serum). Their judgement was thus completely without prejudice. I am keen to see my observations checked independently, and most warmly recommend this "blind" method for the purpose. Even the chief physician may try to draw conclusions about the nature of the serum (unknown to him) that has been used in a particular case: he will be astonished to see how little he is able to do this.....Neither I nor my assistants Dr Reusz, Dr Schwab, Dr Weber, Dr Lube could detect a difference between the two sera. Dr Koennecke thought the old (antitoxin) serum had a certain advantage, while Dr Rehder declared that if he were to fall ill, he would wish to be treated with the new (horse) serum. The views of these two gentlemen thus neutralised each other (Bingel 1918, p 288).

Although Bingel did not mention ‘blinding’ of the patients participating in his study, it is clear from the context that it was a blinded trial in which patients, caregivers, and observers were blinded to the intervention, and he recommended that others should use blinding in replications of his study (Bingel 1918, p 288).

Bingel insisted on using measurable criteria “in order to achieve an objective overall assessment”, which he contrasted with “impressions” from the bedside. His final report was based on an analysis of 471 patients treated with anti-toxin serum and 466 with normal serum (Bingel 1918, p 309). The results were meticulously analysed and presented in detailed tables, as well as in diagrams and illustrative case-reports. No marked differences were detected between the impact of the two sera on the time to shedding of the diphtheritic membranes, or on mortality: there were 47 deaths out of 471 patients (10.0 per cent) given anti-toxin serum, compared with 49 deaths out of 466 patients (10.5 per cent) given normal serum. Nor were differences detected in subgroup analyses in patients who had had tracheotomies or other complications, or after considering the sources of their infections (for example, from within families). Bingel was also aware of the need to study large numbers of patients to reduce the effects of chance, and claimed that his sample had been “sufficiently large to prove that no preference can be claimed for anti-toxin serum” (Bingel 1918, p 331).

Bingel’s challenging results provoked strong and often emotional reactions, which were reported in the lay press. After all, an achievement crowned with a Nobel Prize was at stake, and thus the prestige of German basic research. In general, the paper was simply dismissed. In response to the sceptics (Bingel 1919), Bingel pointed out that he had never contested either the therapeutic or the prophylactic value of Behring’s anti-toxin serum. He acknowledged a valid criticism that the normal serum might have contained some anti-toxin, and admitted that such a possibility had not occurred to him, as he assumed that what he had bought from industry labelled as “normal horse serum” was what had been advertised. To address the criticism he commissioned analyses of samples of the normal serum he had used. These did indeed reveal a very low concentration of anti-toxin in some of the samples tested - roughly 1-3 international units per cubic centimetre. Bingel deemed this to be so low that it could not have had any material effect when compared with an anti-toxin concentration of 500 international units per cubic centimetre, leading to patients receiving an average total dose of between 2000 and 8000 international units (Bingel 1919, p 739).

This argument prompted some of Bingel’s critics to admit that anti-toxin serum must contain other

unspecified therapeutic elements besides the anti-toxin, but it did not stop the defenders of anti-toxin serum continuing to dismiss his provocative findings. As Bingel pointed out later (Bingel 1949, p 101), their views implied that 1-3 international units of serum were as effective as 500 units. He was also criticized for withholding from patients a proven effective therapy in order to test a patho-physiological hypothesis. In response, he asserted that the final decision about the therapeutic value in man of a drug stemming from animal experiments, however well justified theoretically, remained with clinicians, and that it was also clinicians' business, and not that of serologists, whether such studies were to be regarded as consistent with medical ethics (Bingel 1919, p 740).

Bingel's defence was clearly persuasive to some clinicians, however, and his request that his trial be replicated was taken up by some other researchers. In 1933, for example, stimulated by Bingel's findings and their own observations of the effects of normal horse serum on the Schick reaction, Hottinger and Toepfer (1933) reported four separate trials using alternation to generate comparison groups, some using untreated controls. The largest of these trials alternated 400 patients to either anti-toxin serum or normal serum (Hottinger and Töpfer 1933). Like Bingel, Hottinger and Toepfer were unable to detect any differential effects of the two sera, although they stressed that their sample sizes had been inadequate to dismiss the possibility that important differences might exist. They also reiterated that the responses of infected animals to sera might well differ from the responses of diseased patients.

In 1941, more than 25 years after publication of his first controlled trial, Bingel returned to research on serum treatment of diphtheria (Bingel 1949). He explained for his readers that, after the general disapproval of his 1918 paper, he had given in "to the decision of the majority of physicians" and used anti-toxin serum for over twenty years, and he admitted that he had slept better during those years. However, he also admitted that his decision to fall in line with the majority of his colleagues had been made "without inner conviction", and the issue continued to nag him intellectually. He could not agree with those who held that scientific questions that had not been addressed at the beginning of the serum therapy era could not be addressed later: "I am of the opinion that it is never too late to uncover a scientific fact, even if notions that have held up steadily over decades begin thereby to falter". These considerations, and possibly the large difference in price between anti-toxin serum and normal serum, prompted him to take up his comparative studies again in 1941.

In midst of the Second World War, between 1 April 1941 and 30 September 1942, Bingel enrolled nearly 1000 patients in a further comparison of anti-toxin serum and normal serum. The criteria for inclusion were clearly defined, and he reported that allocation to the comparison groups had again been by alternation, a method to which he now sought to give credibility by referring to Paul Martini's support of the approach (Martini 1932; Stoll 2004): "With this 'alternating method' I believe I would achieve statistically irreproachable [*einwandfrei*] results (see M a r t i n i)" (Bingel 1949, p 101). He again used 'cover names' ('blue serum for anti-toxin serum and 'red' serum for normal serum), which suggests that he repeated his earlier double blind methods (Bingel 1949, p 101), although nothing more specific was said about concealing the identity of the two sera being compared. This time, the anti-toxin concentration was checked in the normal serum and found in a concentration of only 1/100th of an international unit per cubic centimetre. As previously, Bingel's assistants and the head nurse, judging from their clinical "impressions", differed over whether they preferred the 'blue' or the 'red' serum. Between 1 October 1942 and 31 December 1943, he ran a third controlled trial, again comprising 1000 cases. The Table summarizes the mortality experiences in Bingel's three controlled trials:

	Normal serum		Anti-toxin serum	
	deaths/N	%	deaths/N	%
1 st trial, 1913-1916	49/466	10.5	47/471	10.0
2 nd trial, 1941-1942	45/458	10.0	57/514	11.1
3 rd trial, 1942-1943	28/492	5.7	32/518	6.2

Bingel emphasises that "Solely the method of alternate treatment in a great number of patients protects from false conclusions" (Bingel 1949, p 103), yet his two later trials do not appear to have been as rigorously conducted as his first trial seems to have been. The differences in the sizes of the comparison groups in the second and third trials suggest that the alternation scheme must have been breached quite often, or alternatively that information on patients was lost in those

times of war. Furthermore, in his report of these trials he reveals that patients treated with normal serum (and by implication, not those allocated anti-toxin serum) received some additional therapies (for example, injections of their own blood). Not only would the comparisons have been confounded because the groups would have differed in ways other than serum type, but maintaining blinding would have presented problems.

Even so, Adolf Bingel's writings reveal considerable methodological and epistemological sophistication, and his application a century ago of a controlled, double blind, clinical trial involving substantial numbers of patients was ahead of its time, not only methodologically, but because his findings posed a serious challenge to the theory which had led to the development of diphtheria antitoxin, and which had been assumed to explain its beneficial effects.

After Bingel had analysed the cases assembled in his three comparative trials, he gave his answer to the question he had posed 30 years earlier. Anti-toxin was not the active agent in serum therapy of human diphtheria; the sera acted non-specifically "as a stimulant activating the defence forces [of the body]". But Bingel's cautious approach continued: at the end of his last paper on the subject he advised his colleagues to continue using anti-toxin serum until his results had been confirmed by others. But would anyone feel able to do a trial comparing antitoxin with plain serum today?

Acknowledgements

I am grateful to Ted Kaptchuk for drawing attention to Bingel's report in *Intentional ignorance: a history of blind assessment and placebo controls in medicine*. *Bull Hist Med* 1998;72:389-433, and to Johanne Sanders-Gieschke for providing some biographical information about Adolf Bingel.

References

Bingel A (1918). Über Behandlung der Diphtherie mit gewöhnlichem Pferdeserum. *Deutsches Archiv für klinische Medizin* 125: 284-332. (Also as a book. Leipzig: Vogel).

Bingel A (1919). Zur Behandlung der Diphtherie mit gewöhnlichem Pferdeserum. *Deutsche Medizinische Wochenschrift* 45:739-740.

Bingel A (1949). Wirkt das Diphtherieheilserum bei der menschlichen Diphtheriekrankheit spezifisch durch seinen Antitoxingehalt oder unspezifisch? *Deutsche Medizinische Wochenschrift* 79:101-103.

Hottinger A, Toepfer D (1933). Über den Wert der Serumtherapie bei Diphtherie, insbesondere bei der malignen, toxischen Form. *Zeitschrift für Kinderheilkunde* 54:505-529.

Martini P (1932). *Methodenlehre der Therapeutischen Untersuchung*. Berlin: Springer.

Russell WT (1943). The epidemiology of diphtheria over the last 40 years. *Medical Research Council Special Report Series No. 247*. London: His Majesty's Stationery Office.

Stoll S (2004). Paul Martini's 'Methodology of therapeutic investigation', 1932. James Lind Library (www.jameslindlibrary.org).

Weindling P (1992). From medical research to clinical practice: serum therapy for diphtheria in the 1890s. In Pickstone JV, ed. *Medical innovations in historical perspective*. New York: St Martin's Press, p 72-83.