

[Download key passages/title pages as a PDF](#)[Download article as a PDF](#)

Bullowa JGM (1928a). Use of antipneumococcic refined serum in lobar pneumonia: data necessary for a comparison between cases treated with serum and cases not so treated, and the importance of a significant control series of cases. JAMA 90: 1354-1358.

Key passages

ALTERNATION

There must be no selection of patients. At Harlem Hospital each alternate patient with pneumonia is placed in the serum series. Because of the structure and organization of our hospital, the male and female cases are conducted as separate units, with separate house staffs, but with the same visiting and resident physicians constantly in charge of all patients with pneumonia. It is not practicable to alternate the cases according to type on admission, as this might occasion a delay of many hours or days. The injection of a powerful polyvalent serum of types I and II, and soon of III also, assures prompt treatment of the cases selected for serum. Only the order of arrival in the ward determines whether a patient is to receive serum.

The delays in typing are occasioned by absence of sputum, by the time requisite for the growth of organisms in the peritoneum of a mouse, or for development of a subculture from its heart's blood. Sometimes a blood culture from the patient, or aspiration of a serous effusion or of the lung itself, furnishes the first information as to type.

In spite of the fact that we were impressed, as the research proceeded, by the beneficial effects of the serum, the question arose as to whether there was conclusive proof that the serum is of value. What test can we apply to our data to see whether proof is adequate and how can we determine the number of cases necessary for a fair evaluation? This leads us into a brief digression concerning what difference in results is statistically significant, and the meaning of the standard error.

If we take a number of samples from a given material (such as a population of hospital patients with pneumonia) and measure these samples for any particular quality (such as case fatality), the measures obtained will vary somewhat, from sample to sample, and if plotted in a diagram will appear as a bell-shaped curve such as

TABLE 7.—Number of Days Elapsed Between Onset of the Disease and Return of Temperature to 100*

Number of Days Elapsed	With Serum		Without Serum	
	Number	Per Cent	Number	Per Cent
Total cases.....	38	100	29	100
Less than five days.....	9	24	2	7
Five to nine days.....	26	68	15	52
Ten days or more.....	3	8	12	41

* Recovered pneumonia patients treated (a) with serum and (b) without serum; type I cases, without bacteremia.

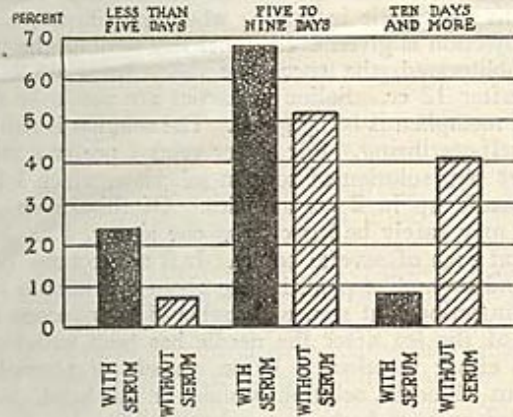


Chart 6.—Number of days elapsed between onset of the disease and return of temperature to 100; type I cases without bacteremia.

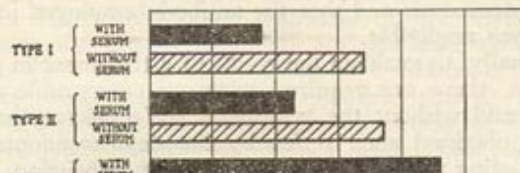
that shown in chart 8. If we do this for two separate materials, such as pneumonia patients treated with and those treated without serum, we shall obtain two separate bell-shaped curves. If the two separate materials are essentially different in their behavior as regards the quality measured, the peaks of the two curves will fall some distance apart, because the measure found most frequently in the one set will differ from the measure found most frequently in the other set. For example, the average case fatality in serum-treated patients is found to be different from the average case fatality in patients treated without serum. But if the difference is small, the peaks of the two curves will fall close together, and if the curves are rather flat, they will overlap and be practically indistinguishable.

The effect of using larger samples in such a case is to contract the two bell-shaped curves, making them high and narrow; the two separate peaks then become more clearly apparent, even though the distance between them remains the same.

This is the reason why it is necessary to have sufficiently large samples, for only thus can one detect with certainty relatively small differences in measurements obtained from two presumably different kinds of material.

TABLE 8.—Deaths per Hundred Patients. (Deaths Within Twenty-Four Hours of Admission Excluded)

Type	(a) With Serum			(b) Without Serum			Difference in Case Fatality (a - b)	Ratio of Difference to Its Error
	Cases	Deaths	Deaths per 100 Cases	Cases	Deaths	Deaths per 100 Cases		
I	55	10	18 ± 5.2	53	18	34 ± 6.5	-16 ± 8.3	1.9
II	26	6	23 ± 8.3	38	14	37 ± 7.8	-14 ± 11.4	1.2
III	21	11	46 ± 10.2	17	7	41 ± 11.9	5 ± 15.7	0.3
IV	54	7	13 ± 4.6	83	7	9 ± 3.2	4 ± 5.6	0.7



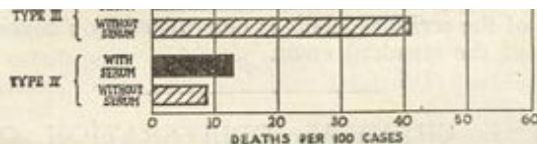


Chart 7.—Deaths per hundred patients (deaths within twenty-four hours of admission excluded).

Without going into technical details, I may explain that the relative spread or flatness of the curve is measured by what is called the standard error of the measurements, which is to the probable error as 3 to 2. In order that the difference between measurements in two separate materials shall be recognizable as definitely significant, the distance between the peaks of the two curves must satisfy a certain statistical test; namely, that the difference between the average measurements in the two cases shall be at least equal to twice the "standard error" of that difference.

In table 8 is a column giving the difference of case fatalities for serum-treated and for non-serum-treated patients. In this column there is appended to each entry a second figure, preceded by the sign \pm . This figure is the "standard error" of the difference to which it is appended.

Finally, there is a column showing the ratio of the difference to its standard error. Whenever this ratio falls below 2, we are not in a position to judge whether any significant meaning is to be attached to the difference; that is to say, whether it is purely accidental or due to a real effect of the treatment. The further this ratio is below 2, the greater must be our doubt.

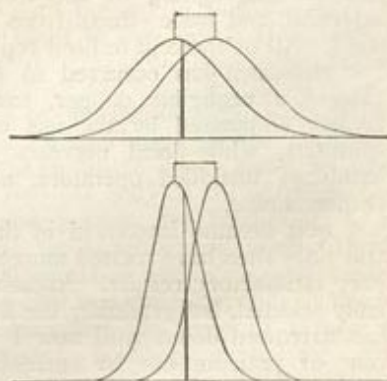


Chart 8.—Influence of size of sample on significance of difference between two rates.

In the type I cases we have practically obtained a result which is twice the standard error, 1.9. A greater difference in the percentage recovery of treated cases than those untreated may be accomplished by future improvements in the serum, and by earlier administration. Even though the serum were to remain as it is, and the difference in the percentage of deaths the same as at present, 16 per cent, a greater number of cases would reduce the standard error and carry conviction of the value of serum treatment. In type II cases the difference is less significant. An effort must be made to make the serum more potent, larger doses may be needed, and more observations of its use must be accumulated. When for type I pneumonia the ratio of the difference in percentage fatality between serum-treated and non-serum-treated cases, to its standard error, becomes more than 2 or 3 it will be our duty to administer serum in all type I cases and to urge its administration on others.

We are aware that type III cases did not have sufficient treatment and that the antibody employed in type IV was negligible.

Finally, to evaluate the result of a treatment in pneumonia, there are required adequate comparable series with and without the treatment. We believe that we have obtained such series by the devices adopted of alternating patients and rating them on admission. The size of the series requisite is determined by a consideration of the standard error.

