

UNCLASSIFIED

AD NUMBER
AD846820
NEW LIMITATION CHANGE
TO Approved for public release, distribution unlimited
FROM Distribution authorized to U.S. Gov't. agencies and their contractors; Critical Technology; JUL 1968. Other requests shall be referred to Commanding Officer, Fort Detrick, Attn: SMUFD-AE-T, Frederick, MD 21701.
AUTHORITY
Biological Defense Research Lab ltr dtd 13 Sep 1971

THIS PAGE IS UNCLASSIFIED

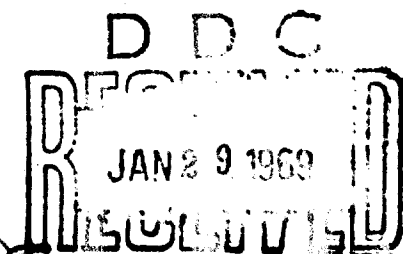
AD846820

TRANSLATION NO. 208

DATE: July 1968

DDC AVAILABILITY NOTICE

This document is subject to special export controls and each transmittal to foreign governments or foreign nationals may be made only with prior approval of Commanding Officer, Fort Detrick, ATTN: SMUFD-AE-T, Frederick, Md. 21701.



Best Available Copy

DEPARTMENT OF THE ARMY
Fort Detrick
Frederick, Maryland

**Best
Available
Copy**

Experimental investigation of chemotherapy in tuberculosis.

by R. Prigge

Klinische Wochenschrift, 20, 633-637 (1941)

Second report (*)

A survey of the literature dealing with the chemotherapeutic susceptibility of tuberculosis to gold compounds a few years ago still presented rather a paradoxical situation: Despite the success which investigators claimed to note in connection with man, it was generally conceded that no exact proof of the efficacy of gold preparations could be obtained in animal tests. In consideration of the fact that the entire field of chemotherapy owes its principles and its development to the animal experiment, attempts were made to overcome this contradiction by theoretical deliberation concerning the mechanism of the gold's effect. It was claimed that the progress of tuberculosis is completely different in artificially infected animals, especially the guinea pig, and in humans; above all it was asserted that the organism of the guinea pig possesses so few inherent defensive potentials that it is helplessly exposed to the progress of the infection and is principally incapable of exploiting the chances given by chemotherapeutic injury to the TB bacilli.

A satisfactory explanation becomes possible only after the investigations of Prigge (1) as well as Rueter and Meyer (2) revealed that experimental tuberculosis of the guinea pig and the white mouse may be influenced very distinctly, so far as the treatment is accomplished only with really appropriate chemical or physical means. After these determinations, the poor suitability of test animals for the demonstration of therapeutic effects could no longer be held responsible for the gold's failure. On the contrary, it suggested itself more strongly to search for the cause of negative results of animal tests in the poor suitability of the gold compounds themselves. The correctness of this explanation is supported by the clinical analyses of Martini et al. mentioned in the first report, in which the ineffectiveness of gold in human tuberculosis, especially pulmonary TB, is exposed unequivocally. We therefore can no longer recognize a principal contradiction between the results of experimental research and clinical tests, and cannot admit of a reason for the neglect of animal-experimental methods for the discovery of chemotherapeutically effective compounds in tuberculosis. In order to compensate for omissions, one should initially deal less with the investigation of the mechanism of chemotherapeutic effect, and more with the development of highly exact techniques which permit a reliable decision on the efficacy of certain groups of substances and the measurement of this efficacy. The extent to which results thus obtained can be fruitfully utilized in the therapy of the tubercular person will, of course, be dependent on the clinician's decision. As Meinschmidt (3) recently stressed with great emphasis: "Although the animal experiment represents the foundation of chemotherapy, there are significant differences between the achievements of chemotherapeutics in the animal test and the effect on human disease."

(*) First report: Klin. Wschr. 1940 II, 1273.

I. Among the methods by which the therapeutic influence on experimentally induced infectious disease, above all on tuberculosis, can be tested, the collective test assumes increased importance. Next to the reaction of the individual animal, it utilizes the total behavior of small or large groups (collectives) of animals as the criterion of effect. Since the principles valid in this connection have been discussed in more detail in previous publications on chemotherapeutic investigations (Frigge 1, 4, 5, 6, 7), more general remarks should suffice here, coupled with a few supplemental data clarifying the progress made in methods.

Every characteristic which allows the collective characterization of an animal group may serve in the evaluation of the chemotherapeutic effect. On the other hand, it appears extremely difficult to reach reliable conclusions solely from the study of pathologic-anatomical alterations. The considerable differences between individual animals which are noted even during undisturbed progress of the infection, make it very difficult to gain a positive evaluation. Investigators who have seen the extraordinarily diverse development of TB among a larger group of identical animals infected in the same manner, and how strongly the findings noted at a certain point differ (even without therapeutic treatment), will consider utmost restraint in the evaluation of anatomical appearances to be indispensable. It becomes especially difficult when comparisons are to be made of animals that have died at different times. An unobjectionable result may be obtained only when all animals belonging to the two groups which are to be compared (they should be as large as possible and comprise the treated series and the control series) are sacrificed at the same time, and when the "stage" of the disease, to be subtracted from the individual tests, is used as a basis for comparison of the two groups. Naturally, the conclusion based on such a procedure will always be somewhat arbitrary — the smaller the compared animal collectives, the more arbitrary. For on one hand we must make the abstraction from individual findings more or less subjectively; and on the other we have no objective method allowing us to evaluate the role chance plays (even in untreated animals) in the distribution of light, medium and severe cases among diverse, apparently identically constituted collectives. Besides, the described method only yields a "cross-section" through the test process by revealing the status of the test at a particular moment, but does not disclose anything about previous or subsequent conditions, while we should be bent on obtaining a "longitudinal section" that shows the entire progression of the test.

II. More reliable, numerically expressible results may be obtained when the course of infection is followed to its end — or to its cure —, in which connection pathologic-anatomical changes found after the test animal's death may serve as an additional criterion. Of the methods used by me for the execution and evaluation of my investigations, I should like to list two that have proved to be especially feasible.

1. Groups of 20 animals (rabbits, guinea pigs, mice), as similar as possible, constituted according to Frigge (17) and Schaeffer's (18) method in such a manner that the different classes of body weight are evenly represented in each series, are infected with human or bovine TB bacilli and in part observed without treatment (control series), in part subjected to treatment (oral, intravenous or subcutaneous) after a few days, generally twice a week. In important cases 2 or more groups, i.e. 40 or more animals, were consolidated

and treated with the same preparation and according to the same test plan, with the same dosage, etc., or were observed as controls. On a suitably chosen "fixed day" the percentage of animals is determined which survive in the different series. If large differences are found, it is determined with mathematical-statistical methods whether they are to be considered characteristic ("significant," conclusive) or the result of chance.

In this connection one should not be misled to assume that the result of comparison increases in certainty when the statistical evaluation of each individual test is conducted on a day on which the survival rates of the treated and untreated animals are most distinctly different. Such a method would even lead to self-deception. A critical evaluation of test results is possible only when the fixed term is set from an objective point of view, independent from the result of individual tests, e.g. so that the result on which evaluation is to be based is determined after the expiration of a certain number of days (from the commencement), or when 50% or 75% of the control animals have died (**), or at another moment chosen as suitable, but in a "non-partisan" way.

This rule naturally should not be applied schematically; above all, it is valid only for tests conducted according to the same program and with the same chemical compound (cf. Section IV). If, for instance, the same course of action is taken with drugs of rapid effect and compounds whose effect becomes active after some time, it would be impossible to grasp the diversity of therapeutic possibilities in a sensible manner. The planning and evaluation of chemotherapeutic tests should not be subjected to a few inflexible rules that allow the convenient use of statistical formulas, but must above all satisfy the multiplicity of biological possibilities. The use of statistical methods therefore represents only an additional process guaranteeing the critical evaluation of numerically expressible test results, but unable to take the place of principles valid for the planning and evaluation of tests.

The result obtainable by the described method may in some cases be extended by repeating the comparison between survival rates on several fixed days; thus it is possible to determine, for instance, whether the therapeutic effect has increased or, conversely, it has been displaced by a dystherapeutic one. This course appears to have its rewards especially when the effective mechanisms of different compounds are to be compared. However, if the survival rate of the treated or untreated animals once attains a lead, no additional conclusions may be reached from the circumstance that the lead is not cancelled out subsequently.

The described method leads to a qualitative result; it answers the general question, whether or not a tested substance is effective. If compounds of ascertained efficacy are available, a measurement of the therapeutic activity may be conducted, provided one of the effective substances is defined as the comparative preparation ("standard preparation"). In this case the control series consisting of untreated animals is replaced by a group of animals treated with the standard preparation. Comparison of survival rates or other values characteristic of the behavior of the animal collection (see below) reveals (***) In especially favorable cases, e.g. in the evaluation of the efficacy of highly effective compounds against pneumococcal infections, the test results are determined when 100% of the animals (control animals) have died.

whether one of the tested substances is more or less effective than the assay preparation, or identical. If the efficacy of different quantities of the tested drug is determined, it can also be found which dose is equivalent to a certain amount of the standard preparation; from this, the ratio of values of the different compounds can be deduced directly.

2. Particularly lucid results may be obtained when the survival ratios found in the treated and untreated animal series are not only compared once or several times, but continuously during the entire test duration and the results are depicted graphically (cf. Fig. 1-4). The graphic representation of the test animals' course of death gives a very clear picture of the total progress of a test and yield the most valuable foundation for the evaluation of the therapeutic efficacy of chemical compounds. Of course, a numerically expressible judgement of the obtained result cannot be deduced therefrom directly. Such a judgement, which should consolidate and evaluate a highly complex, reciprocal effect between macro-organism, micro-organism and chemotherapeuticum extending over a longer period of time, — i.e., a biological process —, in many cases cannot be expressed at all. Still, it may often be useful to get a summary perspective of the test progress by entering each individual animal with a characteristic sign, which significantly influenced the final outcome, in a calculation. Such a sign is the test animals' survival time (or "death time").

Numerous pharmacological, chemotherapeutic, recently even immune-biological investigations have been conducted with the consideration of the survival time; in earlier examinations of the chemotherapeutical susceptibility of tuberculosis I have also utilized this characteristic for the numerical designation of test results (1). Its use is entirely justified, since in many cases it makes the formulation of a sensible average value (and statistical analysis) possible. Still, it has been known for a long time that some experimentally obtained data are very difficult to introduce into the mathematical material yielded by a large collective.

For example, the death times of poisoned or infected animals do not increase freely with decreasing doses of toxin or bacilli. Starting with a certain, more or less precisely determinable limiting dose, we observe that only a part of the test animals die, while the others do not die belatedly, but do not succumb at all to their disease. If the instilled toxic or bacillary dose lags behind a certain "threshold value" (different from animal to animal), the expected effect does not occur with increased retardation, but not at all. This also applies to experimental infection with TB bacilli (cf. Prigge 1, p.445).

The mathematical treatment of such observational data has been attempted in different ways. Kisskalt (8), for example, equates the death time of surviving animals with ∞ . It may be argued that by this method the customary average cannot be arrived at; the arithmetic mean becomes ∞ , as soon as a single animal survives. Other authors ignore the survivors and only account for the dead animals in calculating the average value. Such a course of action must of course appear highly questionable.

The indicated difficulties have been overcome in a very simple manner, one that, to my knowledge, has not yet been utilized. In order to obtain values which allow a sensible calculation of averages even in the cases described, the death time is replaced by the animals' death rate.

If the "death rate" of an animal that dies 1 day (24 hours) after poisoning is designated as 100, the animal dying after $\frac{1}{2}$ day (12 hours) has a death rate of 200, the animal dying after 2 days (48 hours) a death rate of 50, etc. The death rate therefore is defined as the 100-fold magnitude of the reciprocal value of the individual test animal's death time, measured in days. Since surviving animals, designated by ∞ in Misskalt's method, now show a death rate of zero, all test animals can be considered and a reasonable average becomes possible in all cases.

The introduction of death rates leads to the so-called "harmonic mean" of the death times; the latter is the reciprocal value of the arithmetic mean of the death rates defined above, divided by 100. The harmonic mean of the death times can be used in place of the arithmetic mean of the death rate. The construction of both averages is possible in all cases, regardless whether all animals die or one part remains alive. However, since the "harmonic mean" is not customary, it is recommended that only the arithmetic mean of the death times be used in characterizing the death sequence, or else the arithmetic mean of the death rate. I consider the first to be preferable (when both are possible) due to its perspicuity. —

The significance of the survival rates and the mean values is severely limited by the circumstance that the therapeutic influence on an infection is not a uniform process, but a highly complex interplay of powers and contrary powers. Investigations reported earlier support the view that, for example, gold compounds in doses which exclude immediate injury to the organism, may exert an unfavorable influence on the progression of the tubercular infection (Prigge, 1). Observations of the effect of chaulmoogric acid compounds revealed that even therapeutically effective substances may under certain conditions become ineffective or exert dystherapeutic influences (Prigge, 1). Such contrary processes may be covered up in the construction of mean values: The apparent "objectiveness of the number" may in such cases represent hindrances to research. For these reasons a general use of mean values must be avoided. On the basis of years of experience, we consider the graphic representation of death progression to be the most valuable means in the analysis of chemotherapeutic efficacy in experimental tuberculosis, particularly when this method is supplemented by other proven methods (control of weight, examination of the status of the lymphatic glands, pathologic-anatomical examination of all dead animals, treated and untreated, etc.) The value of the graphic method diminishes, however, when applied to small groups of about 5 to 10 animals; the reliability of the graphic method increases with the size of the tested animal collective. It is only in conjunction with this standard system that the evaluation of survival rates and mean values assumes significance. In summing up, their utilization may be characterized as follows: Distinct differences between mean values always represent valuable intelligence, while the absence of such differences does not permit unequivocal evaluation.

III. We may also dispense here with a detailed description of the mathematical-statistical methods used in the critique of experimentally obtained test results, expressed in figures. It is sufficient to point out that the "significance" of the difference between survival rates of treated and untreated animals was tested with the aid of the T-test developed by H. v. Schelling (9), which for the first time guarantees a completely exact treatment of the difference between two percentages, and therefore is logically superior to other systems, even to the χ^2 test of Pearson, which in this case is nearly equivalent numerically, however.

The difference between the arithmetic mean of death times or death rates obtained from two comparative test series cannot be evaluated by heretofore customary methods, since these hinge on the condition that a so-called "normal distribution," i.e. a distribution corresponding to Gauss' law of error, is valid for the observed values, and also because the distribution of death times of tubercular animals (as well as that of death rates) proved to be entirely different from Gauss' distribution of error. The possibility of statistical evaluation of such material has been given only recently; it becomes feasible by means of a method, also developed by H. v. Schelling (10), which is wholly independent of the assumption of a normal distribution.

The utility of the critique of test data, undertaken with the aid of mathematical-statistical methods, becomes apparent primarily by its ability to prevent subjective evaluation of apparently favorable test results, and to give valuable pointers as to which observations would justify further pursuit and the commitment of more extensive investigations. It should be obvious that in this connection the course of action must not be schematic, so that, for instance, only those results are considered which would be designated as "affirmed" by the 3s-method. It seems more important to test in every case the magnitude of probability with which an apparently favorable result may belong to the realm of chance. As long as this value is contained within reasonable bounds, there remains the justification or duty to assume a "systematic" (legal) cause for the observed result and to conduct further observations. The remarks at the close of Section II make it obvious that additional testing of a preparation will appear indicated even in those cases where graphic representation of the disease evokes suspicions of contrary effects, and where the statistical evaluation fails to reveal the presence of efficacy.

IV. In order to avoid disappointments (which would bring unwarranted discredit upon statistics), it should be realized that even the described, exact methods do not make obsolete the necessity for a well-planned preparation of the material for evaluation.

In this paper, which primarily aims at the biologically interested reader, a common error must be pointed out: If of two tests, conducted with the same drug and according to the same plan, one yields a "positive" result, i.e. a result supporting the efficacy of the tested substance, while the other does not show a significant result, it would be wrong to conclude from the judgement obtained by statistical analysis of the first test that the second test should now be discarded as inconclusive. It should rather be considered that the conclusion reached by statistical methods represents an expression of probability, expressed under the explicit condition that no arbitrary selection from

the test results are not true. This does not mean that such tests should be considered, the results of which were obtained by some kind of erroneous process or by the alteration of test procedure or conditions. Only those tests can be considered in composite statistical analysis, which the biologically trained evaluator recognizes as unobjectionable and truly homogeneous. Test results meeting these preliminary conditions should not be isolated, however, but should always be evaluated in context.

The justification for this requirement is easily seen: Even in dealing with completely ineffective chemical compounds, greater differences between the average death times of treated and untreated animals may appear as purely coincidental results, as long as the tests are conducted with sufficient frequency. It is understood that the repetitions of an ineffectual test, in the hope of such a result, cannot yield a result admitting of evaluation. If the hoped-for result should really be achieved after numerous repetitions, — provided the test has been conducted without grounds for objection —, it is easily recognized as the factor so designated in the theory of probability, namely as the "exception" due to chance, which may be expected only in a very large number of cases. Therefore it may under no circumstances be evaluated separately, i.e. by circumventing the other results! On the other hand, it is inconsequential if after various positive test repetitions one result is obtained which — by itself — does not reveal a noteworthy difference between the obtained average values and therefore may not be considered as "statistically affirmed."

If results of several tests, conducted with the same preparation and according to the same plan, are available, every kind of arbitrary selection is prohibited: All available test results must be considered and compiled in the analysis. This is true also of the individual results obtained within the framework of a larger test, e.g. for (apparently disconnected) death times of animals that survived long past the time in which the death of all other test animals had occurred. As long as it cannot be proved that a test error is involved (erroneous deviation in the amount of substance instilled, etc.), it is completely unjustifiable to ignore individual observations which make statistical analysis more difficult. After J. Schaefer (18) (op. cit. p. 62/63) had pointed out that at times even expert statisticians make such errors, our allusions to an unobjectionable test evaluation should be considered more earnestly than may appear necessary to persons not acquainted with the mistakes frequently found in this connection.

V. In the introduction, a concept was illuminated, according to which experimental TB differs from the human primarily owing to the circumstance that the organism of test animals are defenselessly exposed to the disease due to the absence of inherent defenses. Disregarding the question, whether or not an acute infective progress might even be desirable in view of the goal of finding chemotherapeutically feasible substances, (since only highly effective compounds can assert themselves), it must first be admitted that the assumption of a complete or even far-reaching defenselessness of test animals against tubercular infections is false and without experimental support. On the contrary: Frigge's (1) as well as Kuester and Meyer's (2) investigations proved that individual laboratory animals have a distinctly different resistance to tubercular infection, leading to extreme diversity in the behavior of individuals

and the course of disease in a large collective of identically infected animals: next to slight acute tuberculosis we may observe very slowly developing, chronic profections. It may even happen that some animals remain completely healthy and do not even react with glandular tumescence, if the infection's virulence is attenuated. However, it is not only the reaction of individuals belonging to the same collective that is highly variable. The average profection characteristic for the collective, as well as the distinctiveness of the extreme forms, may also fluctuate considerably, depending on the mode of infection. The experimenter may therefore fashion the disease in a manner that appears especially favorable for the execution of the test.

In this connection, as stressed by Huester and Meyer (2), it ought not to be advantageous to change the course of disease by specific alterations of the organism. Some investigators, for instance, have created slowly progressing tuberculosis in rabbits by pre-treating the animals with weakly virulent (human) TB bacilli and reinfection after some time with fully virulent bovine germs. Huester and Meyer point out that in such animals, especially strong natural healing processes set in which make the evaluation of therapeutic agents very difficult, particularly if the investigations deal primarily with the study of pathologic-anatomical changes in the various stages of the disease, the differences of which are to furnish conclusions regarding the disease's tendency to heal.

The exertion of external influence on the progress of disease is not even necessary, as the infection's dose according to quantity and virulence of the bacilli may be considered the method of choice for the creation of diversely calibrated infections, according to extensive experimental experience. If doubts about this fact are expressed time and again, this may stem primarily from the circumstance that the extraordinary scope of dosage variability, which makes a noteworthy calibration of infective progress possible, is unknown to many investigators. We therefore list, in table 1, a synopsis of the differences obtained in the course of tubercular infection of guinea pigs with different doses of the same strain of TB bacilli. In this connection, the question whether the especially high susceptibility of the test animals during the first few days noted by numerous investigators, and the fatal cases not caused by significant anatomical lesions within this period are due to a "tubercular ultravirus" or other factors related to artificial infection, shall be barred from the discussion.

As revealed by table 1, the death times shorten distinctly with increased infective dosages; however, the shortening occurs in an entirely different relation than the increase of the infective dose. A significant acceleration in the average course of infection is obtained only by the increase of the bacillary dose to the thousand-fold quantity — e.g. from 0.1 γ to 0.1 mg and from 0.1 mg to 0.1 g (!). Since the calibration of the bacillary dose generally is set within much narrower bounds, it is not surprising that a correlation between infective dose and infective progress has often been doubted. In this connection it should be remembered that the amount of 0.1 γ of moist TB bacilli, according to customary estimates, still represents the enormous sum of 100,000-200,000 germs.

more important than the strain of injected bacilli in their virulence. While, for the most part, avian bacilli cause only very slowly developing disease, even in guinea pigs, it is possible to cause rapidly progressing TB with some of the certain strains (Tb.3), which predominantly attack the lungs. Under suitable infective conditions, an extremely characteristic, acute disease develops, leading to death within 2-3 weeks (Prigge 1, Maester & Meyer 2). The death times of these disease forms are concentrated on a very narrow interval. Such differences in the progress of disease, dependent on the virulence of the infecting strain, are frequently seen in the course of extensive investigations, although not always in such obvious form.

(Conclusion follows)

Table 1. Death times (in days) of guinea pigs after subcutaneous infection with human TB bacilli, strain Tb.1; infection on 8 March 1938.

Animal #(*)	Infective dosis			
	Series A 0.0001 mg	Series B1 0.1 mg	Series B2 0.1 mg	Series C 100 mg
1	14	10	12	10
2	15	27	13	10
3	23	34	28	20
4	27	36	30	20
5	39	39	42	25
6	41	42	44	32
7	45	43	45	33
8	45	44	45	35
9	53	52	47	38
10	57	52	52	39
11	76	53	52	39
12	76	53	57	46
13	77	53	58	48
14	83	55	65	49
15	98	61	67	49
16	111	66	68	50
17	112	68	68	52
18	114	75	69	52
19	138	76	70	54
20	161	103	71	72
Average:	71	52	51	39

(*) The animals are ordered according to their survival time.