Was it statistically legitimate to combine data from the four textile mills in Brachman *et al.'s* (1962) study of the effectiveness of a human anthrax vaccine?

Walter R. Schumm, PhD
School of Family Studies and Human Services
Justin Hall, Kansas State University
1700 Anderson Avenue
Manhattan, KS 66506-1403 USA
Phone: +1 785 539 3641 (home) +1 785 532 1494 (office)
Fax: +1 785 532 5505
Email: Schumm@ksu.edu (work) WRSchumm@aol.com (home)

Abstract

In late 2003, the Brachman et al. (1960, 1962) field study of an earlier anthrax vaccine became one important basis for an FDA regulatory determination that the currently licensed vaccine is effective against *B. anthracis* strains, regardless of the route of exposure. One issue overlooked earlier (Schumm, Brenneman, Arieli, Mayo-Theus, and Muhammad, 2004) was whether or not it was legitimate, from a statistical perspective, to combine the results from the four textile mills to assess the effectiveness of the anthrax vaccine. Therefore, the Brachman et al. (1962) field study was again reexamined in terms of its statistical validity. The Box's M test, which evaluates the statistical legitimacy of combining data from different groups (in this case, the four mills), was very significant ($p < .001$) in all three statistical tests performed, indicating that the data from the four mills should not be combined. Arguments for combining or pooling the data from the four mills cannot be justified from statistical or scientific evidence.

Keywords: Anthrax vaccine, limitations of research, Box's M test, pooling data, vaccine efficacy

1. Background

The safety and the efficacy of the current anthrax vaccine used by the U.S. military has been challenged [1,2] despite arguments in its favor [3]. The U.S. Food and Drug Administration (FDA) has responsibility to license vaccines based on its scientific assessment of the efficacy and safety of such vaccines. On December 30, 2003 the FDA issued a regulatory opinion on anthrax vaccine, a Final Rule. FDA's Final Rule, and other assessments of the efficacy of the current anthrax vaccine in humans, are primarily tied to the reputed success of the field investigations done, with a similar vaccine, between 1955 and 1959 at four goat hair mills in the northeastern United States. However, in deciding on the case John Doe #1 et al. v. Donald H. Rumsfeld, et al., (U.S. District Court for the District of Columbia, Civil Action No. 03-707) District Court judge Sullivan issued an order on October 27, 2004 that the FDA had not provided adequate review for the anthrax vaccine and suspended its administration to members of the U.S. military until the FDA had time to obtain further comments on the anthrax vaccine.

2. Issue

Many of the limitations of the Brachman et al. [4] design [5] and its statistical errors and limitations [6] have been detailed elsewhere. However, the Food and Drug Administration (FDA) has argued that it was legitimate to combine data from the four textile mills used in the Brachman et al. [4] study of the effectiveness of a human anthrax vaccine among mill workers at four mills (the locations of the mills remain in dispute). While admitting that "the five cases of inhalation anthrax reported in the Brachman study are too few to support an independent statistical analysis" the FDA argued that both inhalational and cutaenous cases [from all four mills in the study] had been combined, thereby making the results applicable to both routes of exposure [Proposed Rule and Proposed Order, 29 Fed. Reg. 78286, December 29, 2004; http://www.fda.gov/cber/rules//bvactox.pdf]. Combining data from the four mills might seem to be desirable because it would increase the statistical power of the dataset. It might also make it appear that the vaccination program was statistically successful at all four mills when in fact it was only significant statistically at one of the four mills [6: 178, Table 11]. However, Schumm et al. [5, 6] failed to test the statistical legitimacy of combining data from all four mills, leaving it as an important issue to be resolved scientifically.

3. Method

The Box's M test, an output from the SPSS discriminant analysis routine [7] is used to determine the statistical legitimacy of combining data from different groups [7: 108-109]. Here, the different groups were the four textile mills used in Brachman et al. [4]. The total sample contained 1,249 cases, whereas the smaller sample of only treatment and control (placebo) subjects included only 793 cases.

The variables used to create the covariance matrices for each of the four groups included risk (low, high), vaccine (none or placebo, partial vaccination, and actual full vaccination), and illness (anthrax infection or not), and partial vaccination (yes, no). In the SPSS discriminant analysis programs used to obtain

doi: 10.1588/medver.2005.02.00042

Box's M test [7], prior probabilities were calculated from the sample size of each of the four groups.

Three tests were performed. Because the experimental group (total N = 793) did not contain data on partial vaccinations, only three variables could be included in each group's covariance matrix. Therefore, three variables were used for both the experimental (N = 793) and the total groups (N = 1,249). For the total group, four variables were used as well; adding the partial vaccination variable (which was weakly correlated with the vaccine variable, r = .11).

4. Results

Using three variables with 1,249 subjects, Box's M test yielded a score of 142.58, with an approximate F test of 7.88 with 18 and 1,394,368 degrees of freedom ($p < .001$). Using four variables, Box's M test yielded a score of 292.65, with an approximate F test of 9.68 with 30 and 1,160,713 degrees of freedom ($p < .001$).

Using three variables with 793 subjects, Box's M test yielded a score of 65.99, with an approximate F test of 3.63 with 18 and 529,086 degrees of freedom ($p < .001$).

Observation of the correlation matrices for the four groups suggested that differences among the four groups were not only statistically significant but were substantial as well, with the vaccine-illness correlation being statistically significant for Mill S but not the other three mills.

5. Discussion

The results clearly rejected the statistically validity of combining data from the four mills. Consequently, it is not appropriate to analyze the data from Brachman et al. [4] for the entire "clinical" trial of the human anthrax vaccine together in one combined group, for either the total group (N = 1,249) or the experimental group (N = 793).

6. Implications

While it may be politically desirable to combine the data from the four textile mills, it is not scientifically appropriate. Given that, it remains clear that the anthrax vaccine demonstrated only cutaneous efficacy at just one mill (Mill S)–and not the mill (Arms Mill, Mill A) where the inhalation cases of anthrax infection occurred. Given that the results for Mill A never were and are not now significant ($p < .05$) even when including the cutaneous cases of anthrax infection, our results here further demonstrate that there never has been any *scientifically valid* basis for arguing that the vaccine was effective against inhalation anthrax in Brachman et al's study [4]. In fact, at only one of the four mills was the vaccine effective against cutaneous anthrax infection. Inferring from Mill S alone that the vaccine was effective against cutaneous anthrax requires an assumption that the data from the four mills can be combined legitimately. Of course, it might seem logical to infer that partial success of the vaccine leaves us with our best hope to prevent both cutaneous and inhalation anthrax infections. However, science rests on proving facts, not merely wishing them to be so.

References

[1] Nicholson GL, Nass M, Nicolson NL. Anthrax vaccine: controversy over safety and efficacy. Antimicrobics and Infectious Disease Newsletter, 2000;18(1):1–6.
[2] Nass M. The anthrax vaccine program: an analysis of the CDC's recommendations for vaccine use. American Journal of Public Health, 2002; 92:715–21.
[3] Friedlander, AM, Pittman, PR, Parker, GW. Anthrax vaccine: evidence for safety and efficacy against inhalational anthrax. JAMA, 1999; 282: 2104–6.
[4] Brachman PS, Gold H, Plotkin SA, Fekety FR, Werrin M, Ingraham NR. Field evaluation of a human anthrax vaccine. American Journal of Public Health, 1962;52:632–45.
[5] Schumm WR, Brenneman RL. How "adequate and well controlled" was the "clinical trial" of a human anthrax vaccine, 1955-1959? Medical Veritas 2004;1(2):166–70.
[6] Schumm WR, Brenneman RL, Arieli B, Mayo-Theus S, Muhammad JA. statistical reanalysis of Brachman et al.'s (1962) study of a human anthrax vaccine. Medical Veritas 2004;1(2):171–8.
[7] Norusis MJ. SPSS-X Advanced Statistics Guide (2nd ed.) Chicago:SPSS, Inc., 1988.

doi: 10.1588/medver.2005.02.00042