Unanswered questions and ethical issues concerning US biodefence research

W R Schumm,1 R R Nazarinia,1 K R Bosch2

ABSTRACT
Unanswered questions and ethical issues associated with US biodefence medical research over the past five decades are discussed. Objective scientific standards are essential for making policy decisions that can stand the test of time. For decades, scholars have reported that the human anthrax vaccine field trials conducted in the 1950s by Brachman and his colleagues were single-blind rather than double-blind. Nevertheless, in March 2005, Dr Philip S Brachman reported in a letter to the US Food and Drug Administration that his study had been double-blind. It is here argued that, rather, the field trial of a human anthrax vaccine should continue to be deemed as single-blind unless more detailed information is provided to explain exactly how the investigators were kept unaware of which subjects were in the treatment and control groups. Moreover, a number of other questions about the details of this critically important study have remained unanswered and are discussed. More recently, similar concerns have arisen with respect to more contemporary biodefence research, especially with reference to the Federal Bureau of Investigation’s allegations that Dr Bruce Ivins, a US government biodefence researcher, was responsible for the anthrax letter attacks of fall 2001. The medical ethics and related issues involved with continuing to base national biodefence and public health policy on unclear, if not contradictory, research are discussed.

Anthrax is a spore-borne disease caused by Bacillus anthracis that can be introduced into the human system through the skin, digestive tract or lungs. The disease originally occurred in herbivores, which can become infected after contact with spores from soil or infected carcasses. The spores are resistant to destruction and can survive for long periods of time in the ground. While fatalities are rare from treated skin or cutaneous anthrax,1 fatality rates are high for the other two types of anthrax infection, as illustrated most recently in the anthrax attacks of autumn 2001, which involved at least five mailed letters resulting in at least 22 confirmed or suspected cases of anthrax infection, 11 of which were inhalational, killing five victims.7 The letters were mailed from a mailbox on Nassau Street in Princeton, New Jersey, but were postmarked from Trenton, New Jersey. In recent centuries, the most common human exposures have occurred either agriculturally or in textile mills that were processing animal hair products contaminated with anthrax. Consequently, the early investigations of anthrax vaccine occurred in textile mill settings. Many have considered anthrax as a viable biological warfare weapon.2–5 Hence, there has been deemed a corresponding need for defensive measures against anthrax, regarding both prevention (vaccines) and treatment (vaccines and/or antibiotics). Work on a vaccine began in the late 1940s and early 1950s, eventually leading to vaccine testing at four textile mills, one in New Hampshire and three in the Philadelphia area, between 1955 and 1959. Work continues currently to improve anthrax and other biodefence vaccines.

EARLY TESTS QUESTIONED
The efficacy and safety of human anthrax vaccine have remained controversial.6–9 An important reason behind the controversy has remained: several critical limitations of the early vaccine tests at the four mills. One limitation has been the issue of blinding in those early tests. The US Food and Drug Administration (FDA) recently promulgated a Final Order10 on a human anthrax vaccine in which it appears to have accepted, or at least not disagreed with, a sentence in a letter from Dr Philip S Brachman11 in which he described his 1950s study12–14 as double-blind, stating specifically, “In the study that I conducted [sic] in the late 1950s evaluating human anthrax vaccine developed by scientists at Fort Detrick, Frederick, Maryland, we selected our volunteer group of subjects, employees at four textile mills in the northeastern part of the United States who volunteered for this double-blind evaluation.” In the same letter, Dr Brachman went on to say, “At the time of my evaluation of reactions, I was not aware as to whether the employee was in the vaccinated or placebo group,” even though he also stated that “I personally gave every dose of vaccine and placebo.” Nevertheless, all interested parties appear to agree on the relative importance of the Brachman studies,11–15 despite their age, regarding the merits of human anthrax vaccine. It appears that the FDA also accepted the credibility of the early Brachman studies.9 Here we are not attempting to propose a speculative origin for a new disease14,15 but rather to insist that, in spite of some historical evidence about the prevalence of legal fictions in medical ethics,16 medical researchers ought to account for discrepancies in their research, especially when their testimony is foundational to national public health policy. The role of the intelligence community in informing potential vaccine recipients about threat levels is also important for risk-benefit decision-making17 but is not considered here.
DOUBLE- OR SINGLE-BLIND?

A double-blind study is defined in medical research as

... an experiment designed to test the effect of a treatment or substance by using groups of experimental and control subjects in which neither the subjects nor the investigators know which treatment or substance is being administered to which group. In a double-blind test of a new drug, the substance may be identified to the investigators by only a code.

Another definition states that:

It is double blind because both the subjects and those in contact with them are blind to details of the experiment.

The acceptance of the “double-blinded” assertion is important because the FDA\(^{20}\) has historically considered double-blindness an important characteristic of high-quality clinical research. At the same time, the FDA has highlighted the importance of scientific consensus about study adequacy—for example, there should be scientific consensus that a study had specific characteristics, not just one person’s opinion on the matter.\(^{29}\) Recent experimental tests on human vaccines have been double-blind,\(^{21}\) for example.

However, the original Brachman study\(^ {11}\) states only that “The employees were not told which material they received,” indicating that the study was single-blind (the researchers were aware of who received vaccine and placebo). Numerous recent professional and governmental sources\(^ {12-21}\) have indicated that the Brachman study was only single-blind, while other recent reports simply fail to discuss the matter\(^ {22}\) or merely acknowledge that the researchers (ie, Brachman and colleagues) failed to specify “whether the investigators were also blinded to the allocation”.\(^ {12}\) Before several hearings of the US Congress, the former Director of the Center for Biologics Evaluation and Research (CBER) of the FDA, Dr Kathryn Zoon, indicated that the studies had been single-blind,\(^ {22-25}\) as did Melinda Plaisier, Associate FDA Commissioner for Legislation, in a subsequent letter to Congressman Burton.\(^ {26}\)

Despite searching hundreds of articles on anthrax vaccine, we ourselves have found no published scientific articles confirming that the Brachman study was double-blind, suggesting, at the very least, a paucity of scientific consensus on that issue. It is particularly notable that Dr Friedlander, who has authored treatises on anthrax with Dr Brachman,\(^ {24,26}\) has described the study by Brachman and colleagues\(^ {11}\) as single-blind.\(^ {24}\) Of course, the original authors of the Brachman studies\(^ {11-13}\) have, for decades, been publishing numerous treatises on anthrax vaccine. However, to our knowledge, not even those authors have ever claimed in any of them that their original study was double-blind. This confused state of affairs leaves open the obvious question, was the study by Brachman and colleagues\(^ {11}\) double-blind or not?

CHANGES IN COMPOSITION OF TREATMENT AND CONTROL GROUPS

As noted, the concerns about blindedness are only one of several. There are many other unanswered questions about the Brachman\(^ {11}\) study. In an earlier paper,\(^ {21}\) the control and treatment groups at the Arms Mill were described as having 150 subjects each, but later\(^ {11}\) they were described as having 149 and 164 subjects each, respectively. Why did the subject counts for a 1957 study change between 1960\(^ {12}\) and 1962?\(^ {11}\) Did someone become “unvaccinated”? Or was record-keeping inadequate with respect to which subjects had or had not received anthrax vaccine?

Similarly, the 1960 study\(^ {12}\) indicated that there were four placebo subjects enrolled in the study at the Arms Mill who contracted anthrax: “Four of the 9 patients in this epidemic had received the vaccine (Table 1). Four of the nine had received the placebo inoculations (no. 3, 4, 5, and 7); 3 had refused inoculations (no. 1, 2, and 6); and 2 were new employees who had not had an opportunity to receive either material (no. 8 and 9).” Yet in the 1962 report,\(^ {11}\) the 7th case from the 1960 study (the 21st case in Table 4 in the 1962 study) is cited as “placebo-incomplete”, with the following narrative description on page 634: “Two “incomplete” placebo individuals developed anthrax: A 62-year-old female [from mill P] who worked in the drawing department developed her lesion eight months after receiving the last inoculation of her initial series, and a 50-year-old female weaver [mill A] developed anthrax three months after the last inoculation of an “incomplete” initial series.” In other words, the criteria for membership in the control group may have changed between 1960 and 1962; what was it, then, in 1955 or 1957? A 50-year-old female weaver was included in the control group in the 1960\(^ {12}\) study but not in the control group in the 1962\(^ {11}\) study.

Furthermore, it appears that employees (cases 22 and 23, Table 4)\(^ {11}\) who started working at the Arms Mill after mid-August 1957 were not included in the study in either the control or treatment groups, since the two “new employees who had not had an opportunity to receive either material” had onset at the end of October and had been employed for 2½ months.\(^ {11}\) How an additional 14 placebo subjects were found between 1960 and 1962 remains unclear, since apparently at least one “complete” placebo was redefined as “incomplete” and since new employees were not being added to the experimental groups after July 1957 (and all employees were vaccinated after the epidemic for safety reasons). In addition, it was noted that “one probable case of anthrax in a placebo inoculated individual did occur that was not reported. This case is not included in the analysis”, suggesting that another placebo subject had been dropped from the study\(^ {11}\) (in other words, not one of the 14 “new” placebo cases reported in the 1962 study). These inconsistencies in counts of subjects in both the treatment and control groups suggest some significant degree of indeterminacy (or poor record-keeping) with respect to which subjects belonged to which of the two groups.

RANDOMISATION PROCESS AND EXPERIMENTAL DESIGN

Brachman\(^ {11}\) claimed that the subjects “were divided into two numerically equal groups according to their length of employment, age, the department in which they were employed, and the specific job performed”. The US Army Medical Command Military Vaccine Agency\(^ {27}\) has claimed that the Brachman study\(^ {11}\) was randomised. How was that division randomised? The 1960 report\(^ {12}\) indicated that the “allocations were made from the employee roster before volunteers were called for.” If so, how did the rates of volunteer response vary among departments, jobs, length of employment and gender? And how did those varying rates influence the initial randomisation process, if any? Was gender used as a stratifying variable?—it was cited as an important variable\(^ {11}\) in previous industrial anthrax research. Brachman classified the high- and low-risk subjects by mill (A, M, P and S) and vaccination status (complete vaccine, complete placebo, incomplete vaccine or incomplete placebo, and refused),\(^ {11}\) but what were the further breakdowns by age, gender and length of employment within...
each of those 40 cells? Is that data still available for independent multivariate analysis? Without answers to these questions, we have no firm idea of how well the randomisation process, if any, worked in creating truly equivalent treatment and control groups. Even setting aside the Arms Mill, the division of vaccinated and placebo subjects was uneven at the other three mills (230 and 250, respectively).

**WHY THE ARMS MILL?**

How was it that the Arms Mill was added into the study in May 1957 when the other three mills were added in Philadelphia? The arms mill had been in the study much earlier (February 1955, May 1955 and May 1956)?

Adding a fourth mill at quite a distance from the other three would have involved a great deal of additional logistical expense for the project. What special considerations prompted the selection of this distant mill? Our investigation suggests that the Arms Mill had the same parent company as one of the mills in the Philadelphia area, which might be part of the explanation for the selection of the Arms Mill. This mill also had experienced a fairly high rate of anthrax infections over the years, but perhaps then it should have included its employees in the study rather than later. In fact, was the anthrax vaccine evaluation at the Arms Mill, if not the entire study, known as Project N at that time or was Project N completely unrelated to this or any other military research on anthrax or anthrax vaccine? Even if Project N was classified at the time, most classified projects can be declassified after 50 years, a time span that has now elapsed since the conclusion of data collection at the Arms Mill in 1957. At the very least, the timing of the “epidemic” is peculiar. From 1941 to 1956, the Arms Mill had only one fatality from 132 cases of cutaneous anthrax and no cases of inhalation anthrax, with five cases of inhalation anthrax and eight cases of cutaneous anthrax in 1957, an unlikely outcome (Fisher’s exact test, p<0.001; odds ratio 32.50, 95% CI 5.6 to 792.6, using one case of inhalation anthrax for 1941–1956). Even the occurrence of so many cases of anthrax within the months of September and October was peculiar, with seven in 1967 compared with 19 for the other years, compared with 115 cases from November to August for the other years and six cases for 1957 (Fisher’s exact test, p<0.005; odds ratio 6.94, 95% CI 2.10 to 22.9, p<0.002).

**VACCINE SCHEDULE**

It has been argued that as of 1955, the definition of fully or completely vaccinated was a schedule of six inoculations—a schedule even recently described as “cumbersome”—followed by annual boosters, but given that definition, none of the subjects at the Arms Mill could have been fully vaccinated, since the vaccinations did not start there until May 1957 and the epidemic of inhalation anthrax struck in August through October of 1957. Most of the experimental group’s subjects at the Arms Mill would have had only three inoculations, since the fourth scheduled inoculation did not occur until 6 months after the start of the vaccination program, which would have been November 1957 at the earliest for the Arms Mill’s research participants. If the six-inoculation definition was truly in effect as of 1957, then none of the results from the Arms Mill should have been included in the overall study of the four mills. Yet, Brachman showed, for the Arms Mill, that the vaccination status (placebo) of a 65 year old female (E.C., case 17) was “complete” as of 2 September 1957, with two more “completes” by 3 October 1957 (cases 18 and 19, not to mention the “fourth placebo” that was redefined as incomplete between 1960 and 1962, as noted previously). Even at the other mills, subjects were deemed completely vaccinated after only three inoculations of either vaccine (case 5, M.S., mill S, age 35, as of 4 September 1955) or placebo (eg, case 2, J.K., mill S, age 36, as of 30 March 1955). Is there publicly available, printed evidence to validate that the six-inoculation schedule had been formally established as early as 1955 when the research project began, or was that particular schedule a later revision? In other words, exactly when was the six-inoculation schedule first established and reported in the scientific literature? Without a consistent definition of any vaccine’s schedule for both the patient and with respect to how it is used in evaluation studies, it is difficult, if not impossible, to assess its true effectiveness. At the very least, it seems doubtful that the study by Brachman and colleagues actually evaluated the full six-inoculation schedule currently in use, even though the Military Vaccine Agency has argued as recently as late 2005 that the 92.5% effectiveness rate for the anthrax vaccine was obtained by “including only those who were completely vaccinated” with, presumably, what they described earlier as the official “6-dose immunization schedule”. How that can be true when the Arms Mill subjects (who only had three inoculations) were included in the calculations for the 92.5% effectiveness rate is confusing at best.

**IMPLICATIONS**

We think that the unanswered questions and apparent inconsistencies in the design and conduct of these publicly published studies create uncertainties about the procedures followed during and after the research. Since at least two of the principal researchers (Dr Brachman and Dr Plotkin) are still active professionals, these questions and issues could still be resolved for the benefit of the entire scientific community, as well as all others with an interest in biological warfare and public health, as well as the history of medicine in general and of vaccines in particular. To us it seems unethical to allow national public health policy to be founded upon seminal research that remains, to this day, full of contradictions and apparent errors, which could, if the original researchers wished, be clarified to the satisfaction of most medical and public health professionals. We should not be alone in these concerns. Atici and Erdemir recently noted that failures to define subject inclusion and exclusion criteria, to have an appropriate (ie, randomised) control group for true comparison, to have important clinical features distributed equally between the study and control groups, to carry out observations objectively (and accurately report them) or to comment on subjects removed from the study or whose records were lost are important. Such failures, they noted, “affect research outcomes, diminishing the validity and reliability of the research and thus making it scientifically misleading. The unwitting use of biostatistical methods together with conduct contrary to scientific ethics will lead to misinterpretations and partial or false results, thus violating the principles of medical ethics of beneficence, non-maleficence and justice.”

Some might argue that such mistakes and confusion with respect to biodefence vaccines such as anthrax vaccine are a thing of the past. Yet even 50 years later, we remain uncertain whether the Brachman studies were single- or double-blind. We remain uncertain about how many subjects were involved in each group of subjects. We remain uncertain about randomisation. We remain uncertain about the number of inoculations in the full series. We remain uncertain about the nature of Project N and its ties, if any, to the Brachman studies. Still, today,
sources remain in error about the locations of the mills; for example, at least one recent key report indicated that all four mills were in New Hampshire.

CURRENT EVENTS
Similar problems have occurred, at least regarding record-keeping or accurate reporting, in connection with another biovaccine, that for smallpox; the authors of an article on the safety of smallpox vaccine acknowledged numerous errors, especially with respect to sample sizes, but only after their report had been publicly critiqued. The apparent carelessness with reporting details of research in recent smallpox vaccine research cannot but help remind one of the similar carelessness associated with the early Brachman studies.

Likewise, four mill workers died in conjunction with military work on the anthrax vaccine in 1957 under circumstances that were not fully explained at the time nor have been since (ie, the uncertain source of the anthrax spores and why all ill workers were not fully explained at the time nor have been since (ie, the associated with the early Brachman studies.

Researchers cannot but help reflect on the carelessness with reporting details of research in recent smallpox vaccine research. It is not beyond the realm of possibility that on the credibility of scientists working for the government on related issues. It is not beyond the realm of possibility that word of that concern might have reached fellow scientists at Fort Detrick rather quickly, as the memorandum was copied to the Office of the Secretary of Defense (White House Section) on 27 April and was not classified or marked in any way as sensitive material (which otherwise might have required special protected handling). Such background information highlights the importance of a quote from Ivins’s paper: “The only reason to develop a new vaccine is to protect against disease arising as a result of the intentional release of Bacillus anthracis spores by a bioterrorist or in warfare, because the incidence of human disease, particularly inhalational anthrax, is extraordinarily low.” That statement was reiterated in the lead sentence of the article’s abstract on page 55: “The only impetus for the development of new anthrax vaccines is to protect humans against the intentional use of Bacillus anthracis as a bioterrorist or warfare agent.” With Ross Perot, among others, challenging the anthrax vaccine program and the credentials of leading government scientists, Ivins may well have been motivated by a desire to justify continued development of the anthrax vaccine, which might, in his mind, have been both a patriotic outcome and an outcome enhancing the individual and the organisation (Fort Detrick). Given the key Revolutionary War victories achieved at Trenton (1776) and Princeton (1777), New Jersey, the use of these two locations for the anthrax letter mailings may have been an oblique patriotic reference. After the anthrax letter attacks, anthrax vaccine regained its political, if not scientific, acceptance and the annual budget for biodefence increased to over 5 billion dollars a year. Thus, the 2001 anthrax letter attacks may well have achieved their intended objective.

CONCLUSIONS
Clearly, in our opinion, such failures, with which the Brachman studies—and even more contemporary biovaccine studies—seem to be replete, are not a suitable foundation for contemporary national medical or public health policy. In our opinion, there is no excuse ethically for the remaining confusion over the early Brachman studies, the delayed correction of flawed smallpox vaccine research, numerous unexplained civilian deaths or the requirement of 7 years of investigation and millions of dollars (by the FBI) to identify merely a possible source of the 2001 anthrax letter attacks.

In addition to our concerns, others, including the Institute of Medicine, have already noted many failures of the Bush administration’s smallpox vaccination program, which “coincidentally” began and ended with the buildup and end of the initial phase of the Iraq War, with “no apparent public health reasoning behind the decision to offer the vaccine to the public,” resulting in damage to the credibility of both the Centers for Disease Control and Prevention and the government. Another example is the extended delay in publication of the RAND Corporation’s volume on immunizations and Gulf War illnesses, a report languishing since 2002, possibly because its findings concerning anthrax vaccine have not been welcomed by the Bush administration (we will have to wait and see what the policy becomes for the new Obama administration). Perhaps politics had a corrupting influence upon science in the biodefence arena and subjugated truth, ethical practices and justice to a subordinate position. What may not seem controversial to most medical scientists—that medical researchers ought to tell the truth and account for discrepancies in their previous research (not to mention doing no harm)—appears to have been at least a controversial and possibly a secondary consideration within at least some (hopefully isolated) parts of the US biodefence community for a span of as much as five decades. It is our hope that scholarly discussion of these concerns may encourage a re prioritisation, if and as necessary, of the importance of traditional medical and scientific ethics within the US biodefence community.

Funding: No financial support was received from any source to support the preparation of this manuscript.

Competing interests: Dr Walter Schumm provided expert testimony for the plaintiffs in the suing the lawsuit concerning the US Department of Defense directives making the anthrax vaccine mandatory prior to proper review and issuance of a Final Order by the US Food and Drug Administration.

Provenance and peer review: Not commissioned; externally peer reviewed.

REFERENCES
Controversy

7. Jefferson T. Bioterrorism and compulsory vaccination: better vaccines are needed if vaccination is to be made compulsory. BMJ 2004;329:524–5.


42. Schumm WR. Neurologic adverse events associated with smallpox vaccination in the United States: response and comment on reporting of headaches as adverse events after smallpox vaccination among military and civilian personnel. BMC Med 2006;4:27.


