Statistical Analysis Linking U.S. AIDS Outbreak to Hepatitis Experiments

Version 4.2

Author: Thomas R. Keske

TABLE OF CONTENTS

1 Abstract

This statistical study concerns what is probably one of the most significant and overlooked issues of our time. It demonstrates proof of a strong link between the U.S. outbreak of AIDS, and hepatitis studies that were performed on gay males, starting in the late 1970s. The analysis refutes explanations that attribute the connection simply to sexual risk behavior on the part of the study participants.

The analysis also presents evidence suggesting that HIV infections occurring in the studies were more likely to have been intentional rather than accidental. This raises the question of whether the men in these studies might have been used as guinea pigs for covert experimentation, or whether a sexually-transmitted epidemic might have been deliberately induced, as a means to rid society of "undesirables". Regardless of whether the virus itself came into existence naturally, its initial spread was clearly unnatural.

The methodology used in this document is highly similar to that which is typically used to evaluate the effectiveness and safety of vaccines. The analysis evaluates differences in infection rates between a suitable control group, versus a vaccine test group.

In the first two years of the epidemic in San Francisco, between 50 and 60 percent of the earliest known AIDS cases were from persons involved in the hepatitis studies. A goal of this analysis is to calculate specific probabilities for these and other similar figures. It demonstrates that such figures cannot credibly be attributed merely to chance, or to differences in risk behaviors.

Odds of the disproportionate levels of HIV infection among men in the vaccine trial, relative to other men of similar risk behaviors, are shown to be as little as 1 in a trillion.

A statistical link exists not only to experimental vaccines, but also simply to the fact of participation in the hepatitis studies, such as simply to have blood drawn for purposes of monitoring hepatitis prevalence. Few logical or benign possibilities exist to explain why there should be such a connection, yet it exists. Odds against the higher initial rate of AIDS among study participants was as little as 1 in 300,000, when compared to men of equal or higher risk.

Various epidemiological anomalies also suggest that an artificial, simultaneous, mass-infection would have been necessary in order to produce the type of explosion in HIV that was observed in the early 1980s. Full-blown AIDS should have been evident many years earlier, before HIV was nearly so widespread Thousands of infections would have been necessary to fuel the levels of HIV growth that were observed, during years in which no retroactive evidence of HIV exists.

These anomalies are analyzed using computer modeling software.

2 Introduction

Recently, there has been renewed interest as to whether massive polio and smallpox vaccine programs in Africa may have initiated, or at least accelerated, the spread of AIDS on that continent [1]. The possibility that a similar vaccine phenomenon may have occurred in the United States should heighten concern.

For two decades, there has been some concern about a possible connection between the outbreak of AIDS in America and a government-sponsored experimental hepatitis B vaccine which was injected into gay men in New York City, San Francisco, Los Angeles, Chicago, St. Louis, and Denver, between the years 1978-1982.

Several reports in the medical literature attest to the safety of the experimental vaccine. Thus, the connection between the gay experiments and AIDS has been largely dismissed as unworthy of investigation. There is a possible element of bias in this hasty dismissal. The question of possible vaccine contamination is too important to allow political concerns or emotional reactions to interfere with an objective investigation and discussion.

The purpose of this document is to reinvestigate the connection of the original AIDS outbreak in San Francisco gay men not only with the experimental hepatitis vaccine, but also simply with the act of participation in the government-sponsored hepatitis study. This is a new analysis of data collected from various published studies.

This document will demonstrate how gay men who volunteered for government hepatitis experiments were far more likely to become infected with HIV than those who did not take part in such experiments, to a degree not credibly explainable by chance or by life-style.

This document does not speculate whether HIV is an old or a new virus, nor does it explore the outbreak in Africa. It does not focus on whether HIV is a natural virus or a genetically-engineered virus that could have resulted from laboratory experimentation.

None of these scenarios preclude the possibility of HIV contamination of the hepatitis vaccines.

The analysis will attempt to show that there is a significant statistical correlation between vaccine volunteers and HIV infection, that is ***not*** merely the result of “high-risk” sexual behaviors.

This document is being distributed to AIDS activists, virologists, biostaticians, journalists and others, in an effort to promote further research and dialog into the proposed connection of AIDS to the government hepatitis studies.

This document is intended for a broad target audience, including persons of varying backgrounds and levels of knowledge about statistics. For those who have more questions about the statistical computation, a primer and a more extended discussion is provided in Appendix F. Persons who have background in statistics can simply skip this section.

As will be explained, there is good reason to believe that HIV was in the vaccines. There is also reason to believe that the presence of HIV was not likely to have been accidental.

Furthermore, there is an even more peculiar connection of HIV infection simply with participation in the government studies, even for gays who received no experimental vaccine. This connection is not credibly explained simply by the "risk" status of the men involved. This connection is even more disturbing, because there are no vaccines involved for which a possible accident could have occurred in the production.

During World War II, a U.S. State Department official once dismissed allegations about the Holocaust as being of too "fantastic" of a nature to be worthy of forwarding. It may be fashionable in current times to make caricatures of every allegation concerning cover-ups, conspiracies, or secret experiments. However, this fashion does not represent wisdom today, any more than it did in World War II.

There are already more than 33 million people estimated to be infected with HIV/AIDS.

It is perhaps the single most significant incident that has occurred in human history. The number of lives claimed can be expected to exceed the 6 million killed in the Holocaust, or even to exceed the total global battle deaths of World War II. If there is any chance of human agency involved in the genesis of the AIDS epidemic, it is perhaps the most important question of our time.

If there is even the slightest chance of negligence or malfeasance, it would deserve to be investigated. This document will put a quantitative number on that chance, and show that it is more than slight.

The purpose of raising this question is not simply to cast accusation or blame. If a vaccine accident occurred, it is important to determine why, so that such accidents do not happen again. If there is any chance that vaccines could have been contaminated through carelessness or intention, then there must be accountability. If there is a plausible chance that there were further factors causing HIV infection, in addition to the use of vaccines, then those factors must also be identified.

3 HIV Correlation to Vaccine Trial

3.1 Methodology Overview

The primary focus for the statistical analysis of vaccine involvement is based on two distinct groups (cohorts) of San Francisco gay men. One group received the experimental hepatitis B vaccine; the other did not.

In 1978, a research group from the San Francisco Department of Public Health began epidemiological studies of gay and bisexual men attending the City Clinic, a public health clinic for treatment of sexually transmitted diseases.

The San Francisco City Clinic Cohort Study (SFCCC) involved over 6700 gay men. Most of these men participated by donating blood samples for hepatitis study, but did not receive experimental vaccines.

A smaller cohort of 359 homosexual and bisexual men, selected from the larger group of 6700+ men in the San Francisco City Clinic, participated in a clinical trial of a vaccine to prevent Hepatitis B [2]. In this document, this group is known as the San Francisco Hepatitis B Vaccine Cohort Study (SFHBVCS).

The second group referenced in this document for purposes of retrospectively analyzing HIV incidence is the San Francisco Men's Health Study (SFMHS), a cohort of 799 homosexual and bisexual men sampled from 19 high-risk census tracts in San Francisco [2]. The SFMHS began in June 1984. It also included 204 HIV-negative heterosexual men, who are not relevant to this analysis.

Different studies sometimes refer to slightly differing numbers of men in the SFCCC and SFMHS, depending on the date of the study. For example, some cite 6875 for SFCCC and 809 for SFMHS. This document will use values as cited in the context of specific, referenced studies, or will otherwise prefer the higher values.

The SFHBVCS group is composed of 359 San Francisco gay men who received the experimental vaccine. The SFMHS group consists of the 799 high-risk San Francisco gay/bisexual men who did ***not*** receive the vaccine. Both groups were at equal risk of acquiring HIV. Further details on this point are addressed in section 9.

If "equal risk" can be adequately demonstrated, then any differences in the HIV rates between the two groups should be attributable merely to random chance. The analysis will demonstrate that the hypothesis of "random chance" can be rejected.

The statistical comparison of the two groups is a very commonplace type of problem, that is easily computed.

3.2 Lemp 1990 Study

One of the sources of data for this analysis is a study headed by Dr. George Lemp [3]: "Projections of AIDS morbidity and mortality in San Francisco".

Dr. Lemp, formerly with the AIDS Office of The City's Department of Public Health, now serves as director of the University-wide AIDS Research Program at the University of California (starting 1997) [4]

The purpose of Lemp's study was not to evaluate the hepatitis vaccine or to compare the men in the trial with men who did not receive the experimental vaccine. Its purpose was to develop a model for predicting the growth over time of the AIDS epidemic. These projections required tracking of HIV seroconversion in high-risk men. Stored blood samples, taken from the men in the San Francisco Clinic studies, were useful for this purpose. Testing of the blood samples determined exactly when the men showed their first indications of exposure to HIV.

For these purposes of his own, it happened that Lemp collected data which compares the SFHBVCS vaccine group with the SFMHS non-vaccine group. This same data is also useful for the further analysis in this document.

This document does not claim any endorsement from Dr. Lemp. It merely uses the Lemp data for a different purpose. I contacted Dr. Lemp to verify that his data was correct as quoted (see Appendix B).

The data in question, which I verified with Dr. Lemp, was derived from charts contained in the study. It was previously posted to sci.med.aids by Billie Goldberg, a San Francisco lay scientist and AIDS researcher:

SFHBVCS: 1978 - 0.3%, 1979 - 4%, 1980 - 15%, 1981- 28%, 1982 - 40%,
1983 - 46%, 1984 - 47%, 1985 - 48%, 1986 - 48%, 1987 - 49.3%

SFMHS: 1978 - 0%, 1979 - 2%, 1980 - 4%, 1981- 10%, 1982 - 23%,
1983 - 42%, 1984 - 48%, 1985 - 49%, 1986 - 49.3%, 1987 - 49.3%

The above two lines show the total percent of HIV infection, in each year,

for the SFHBVCS (vaccine) group, and the SFMHS (high-risk gay men who did not receive vaccine).

The SFHBVCS group had blood samples taken in each year shown, used to estimate HIV prevalence. The SFMHS had an estimate derived from a subgroup in 1982, and complete samples from 1984, on. The Lemp study used curve-fitting to fill in values for years that did not have actual samples. The analysis in this document will focus primarily on the year 1982, since that is the first figure based on actual measurement, and is therefore likely to be the most accurate.

The rate of HIV infection in the SFMHS group will be taken as a measure of the "expected" rate of HIV for the high-risk population of San Francisco, for this analysis.

The justification for this is based on the comparative patterns of HIV growth in the SFMHS and the SFHBVCS, as shown in the Lemp data.

Both groups start out nearly the same. When each group hit a level of 45+ percent infection, it abruptly ceased the high rate of growth.

Even though the SFHBVCS vaccine group had a slight "head start" in infection, the SFMHS group "caught up" and actually hit the wall of saturation at 49.3% infection, in 1986, a year ***ahead*** of SFCSS vaccine group. This suggests that on average, the SFMHS, non-vaccine group may have been even ***more*** promiscuous and high-risk than their vaccine counterparts.

What most distinguishes the two groups is the large, ***initial*** level of HIV seroconversion in the vaccine group, shortly after they received the vaccine. Based on the fact that SFMHS appears to be at least at equal risk for HIV infection, the two groups should have been roughly equal in HIV prevalence during the earliest years, as well. What the statistical analysis examines is the likelihood for this initial deviation.

3.3 Estimation of High Risk Population

In order to compute probabilities for the vaccine group to have exhibited their high rate of HIV infection by random chance, it is necessary to have an estimate of the total high-risk, gay male population in San Francisco, in 1978. It is important to note that this is an estimate of only the ***high-risk*** population, not the total gay male population. The intent is to eliminate the high-risk status of the vaccine group as the postulated reason for their exhibiting a higher rate of HIV infection.

As it happens, the conclusion of the statistical analysis is ***not*** highly dependent on having more than a rough count of the "high risk" population. This question is analyzed further in section C.1.

The estimate used is based on statistics for the city of San Francisco, for total HIV infections that occurred between 1978 and 1999. It is reasonable to imagine that the numbers of gay men who did in fact become HIV+ in San Francisco, in the two decades to follow, would roughly reflect the numbers who were at risk in 1978 (not accounting for immigration and emigration).

Following is data from the city of San Francisco [5].

```
Reported AIDS cases since 1981:  26,398

AIDS deaths to date since 1981:  18,066

Persons currently living with AIDS: 8,332

Estimated HIV Infected to date (since 1981): 15,250
  (approx. 1 in every 50 San Franciscans (2.1%)


AIDS BY CATEGORY:

   SAN FRANCISCO: 79% MSM, 11% MSM+drug,  7% iv-drug, 1% heterosexual
   CALIF:         71% MSM,  9% MSM+drug, 10% iv-drug, 4% heterosexual
   US:            49% MSM,  6% MSM+drug, 25% iv-drug, 9% heterosexual

   ("MSM" = Men having sex with Men")
```

The total HIV infections in San Francisco is the total number of cumulative AIDS cases, (26398) plus the total number of HIV infections that are not yet progressed to AIDS (15250).

In San Francisco, 79% of the AIDS/HIV cases are gay men (MSM).

An additional 11% is combined "MSM plus IV drug" category. For the sake of this estimate, half of the 11% will be counted as attributable to MSM (this is a small factor, anyway).

The total percent attributed to MSM is then .79 + (.5 * .11) = .84

The total estimate of high-risk gay men is then
(26398+15250) * (.84), or roughly 35000.

A factor that could possibly tend to make this estimate too high is that the estimate does not account for possible immigration into the city. Also, people who acquired HIV over a longer period of time may not have been as high-risk as the people who acquired AIDS in the early years.

However, there are also significant factors that could tend to make the estimate too low. It does not account for men who are HIV+, but simply have not been tested and counted. It does not account for men who adopted "safe sex" practices after the AIDS epidemic became publicized. This would tend to significantly reduce the risk level of men who were previously "high risk."

Thus, some factors would tend to make the estimate err on the side of being too high, while other factors would make it err on the side of being too low. This mix of factors has some tendency to cancel each other out, in overall effect.

The estimate of 35000 high-risk gay men was made independently of the Lemp study. Interestingly, the Lemp study needed to make an estimate of the ***entire*** gay male population of San Francisco. At a minimum, the estimate of 35000 "high risk" gay men should be less than the estimated total of gay males in the city.

In a random phone surveys, Lemp's study estimated 69,122 openly gay males. In another, more conservative estimate, based on extrapolating SFMHS rates of AIDS to the entire city, produced a figure of 42,509. Lemp compromised on a middle figure of 55,816 gay males in San Francisco.

Thus, the 35000 figure for "high risk" seems reasonable. Note that Lemp's phone estimate of 69000+ gay males does not account for persons who would refuse to discuss their sexual orientation over the phone, which may have been significant.

A case could be made for counting ***most*** of the gay males in the city as being "high risk". Consider, for example two men who play Russian roulette, one using 4 of 6 empty chambers on each play, and another who uses only 2 of 6 empty chambers. One has twice the risk of death, per trial. After 25 trials, the "lower risk" man has a 99.996% chance of being killed, versus a virtual 100% for the "higher risk" man.

A similar phenomenon appears to apply to the lion's share of men in San Francisco- within less than a decade, infection rates were nearly 50% in many areas. Thus, the distinction between "high risk" and "low risk" among gay males may be a false dichotomy.

As it turns out, the accuracy of the figure for "high risk" gay men is somewhat of a moot point. Later analysis will show that the net result of the statistical comparison is relatively insensitive to the effect of varying the estimated number of high risk men (Section C.1)

3.4 Vaccine Correlation Analysis

Referring again to the Lemp study data:

```
SFHBVCS: 1978 - 0.3%, 1979 - 4%, 1980 - 15%, 1981- 28%, 1982 - 40%,
1983 - 46%, 1984 - 47%, 1985 - 48%, 1986 - 48%, 1987 - 49.3%

SFMHS: 1978 - 0%, 1979 - 2%, 1980 - 4%, 1981- 10%, 1982 - 23%,
1983 - 42%, 1984 - 48%, 1985 - 49%, 1986 - 49.3%, 1987 - 49.3%
```

Both groups start out with zero or near-zero HIV exposure in 1978. HIV infection exploded in both groups from about 1980, onward.

By 1987, both groups are nearly identical once again, with a nearly 50% infection level. New growth in both groups has slowed dramatically, to little or no new increase.

The behavior is like an “explosion”, starting from next to nothing, spreading very quickly, and infecting a sizeable percentage of both groups. It is not highly meaningful to examine either the very late period (1983-1987), when both groups reached a near-saturation, or to examine the very beginning (1979-1980), when hardly anyone was infected.

The early period from about 1980-1982 shows a more tell-tale difference. In 1982, some 40% of the SFHBVCS shows HIV infection, while only 23% of the SFMHS shows infection.

My question is: What is the exact statistical probability of this difference occurring by “random chance” alone? This is the essence of the computation to follow.

We have estimated a pool of 35000, equally high-risk gay men. From these, we play “God” and draw a random sample of 23% (using the SFMHS infection rate), designating these as the men who will have become HIV+ by 1980. This equals (23% x 35000) = 8050 men. Of these 8050, some (40% x 359) = 144 are from the SFHBVCS vaccine group (the actual result for that group).

What we would have normally expected, on average, would have been (23% x 359) = 83 men, rather than our 144. The probability for this difference between the expected and actual results, by random chance alone, is computed by a standard formula:

Let:

T = total population size (=35000)

v = vaccine subgroup size (=359)

s = sample size - no. of HIV infected men drawn at random from total (=8050)

n = no. of men from the vaccine group who were found in that sample (=144)

The odds for drawing ***exactly*** "n" men would be given by the formula

(vCn * (T-v)C(s-n)) / sCT

where the notation "xCy" is the computation for total combinations in choosing "y objects from a group of x objects":

xCy = x!/((x-y)!y!)

where "!" is "factorial". E.g. x! = x * (x-1) * (x-2) * (x-3) ... * 2 * 1

The odds for drawing "n" ***or more*** is computed iteratively. After computing the odds for "n", we then compute the odds for n+1, n+2, ... s, and sum these probabilities.

This computation was performed using a computer program written by the author, called "comb.c". The program source is not listed here, due to length, but is available upon request from the author. It is written in C programming language, for Unix or for Microsoft Visual C++.

The program output for this problem is as follows:.

Subgroup size = 359
Total group size = 35000
Sample size = 8050
n = 144
PROBABILITY IS: 2.72128e-13

This is 2.7 times 10 to the -13th power, which is a unimaginably small probability
(roughly 1 in 3,700,000,000,000).

The methodology for the above calculation is very analogous to what the researchers themselves used in order to prove that their vaccine prevented hepatitis (explained further in Appendix A). The validity of the logical and mathematical approach is easily demonstrated.

Accounting for margins of error does not alter the conclusion, as discussed in Appendix C.

3.5 Lemp Study, Odds Per Year

Below are computations of the probabilities for all of the years from 1979-1982, that the higher proportion of HIV infection in the vaccine group might have been random chance.

In the earliest year of 1979, at the beginning of the vaccine trial, there is not a statistically significant difference between the vaccine and non-vaccine group. All of the other years show a significant difference.

- **1979: SFMHS =2% =700 ; SFHBVCS =4% =14**

Subgroup size = 359
Total group size = 35000
Sample size = 700
n = 14
PROBABILITY IS: 0.014146 (1 in 71)

- **1980: SFMHS =4% =1400 , SFHBVCS =15% =54**

Subgroup size = 359
Total group size = 35000
Sample size = 1400
n = 54
PROBABILITY IS: 5.62915e-17 (1 in 18 quadrillion)

- **1981: SFMHS =10% =3500 , SFHBVCS =28% = 101**

Subgroup size = 359
Total group size = 35000
Sample size = 3500
n = 101
PROBABILITY IS: 2.2762e-22 (1 in 4 billion-trillion)

- **1982: SFMHS =23% =8050 , SFHBVCS =40% = 144**

PROBABILITY IS: 2.72128e-13 (this is the previous example)

4 HIV Correlation to Non-vaccine Participation

4.1 Over-representation in earliest AIDS cases

Further analysis shows that there is peculiar, unexplained correlation of early AIDS infection in the entire San Francisco City Clinic Cohort (SFCCC). These 6875 men were studied by having blood drawn for purposes of determining prevalence and transmissibility of hepatitis B. Only a much smaller subgroup of 359 men, drawn from entire SFCCC, had received experimental vaccine.

A number of studies suggest that the SFCCC was over-represented among the earliest of AIDS cases, compared to men of equal risk. These studies do not identify if the early AIDS cases from the SFCCC also happened to be the same men who had received the experimental vaccine. If they had in fact received the vaccine, then the failure to mention this would have been a gross omission on the part of the AIDS researchers.

If all or some of these disproportionate SFCCC AIDS cases were actually attributable to the subset of men in the vaccine trial, then it suggests an ***extremely*** strong link to the vaccines.

If the AIDS cases were merely men in the SFCCC who did not receive the vaccine, then it raises the disturbing question of what other unknown factor could explain the disproportionate HIV infection.

There is little in the study that should by rights have put men at significantly higher risk for getting HIV. It is unlikely, for example, that unsterile or reused needles would have been involved in drawing blood.

It must be demonstrated that the men in the SFCCC did not have a higher rate of early HIV infection simply because of such an obvious factor as higher-risk behavior.

Consider how the shape of a graph should appear when a group at higher risk for AIDS infection is compared to a lower risk group:

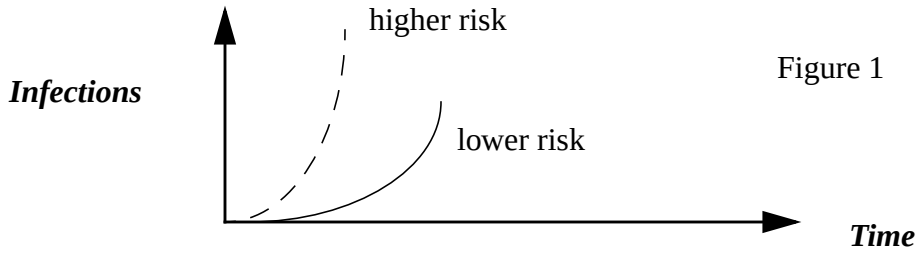

Figure 1

The higher-risk group will tend to produce AIDS cases earlier. The numbers of cases will grow faster, thus producing a sharper, faster rising curve.

Below is a graph showing the growth of clinical AIDS in the SFCCC, versus the entire city of San Francisco, in the first years of the epidemic:

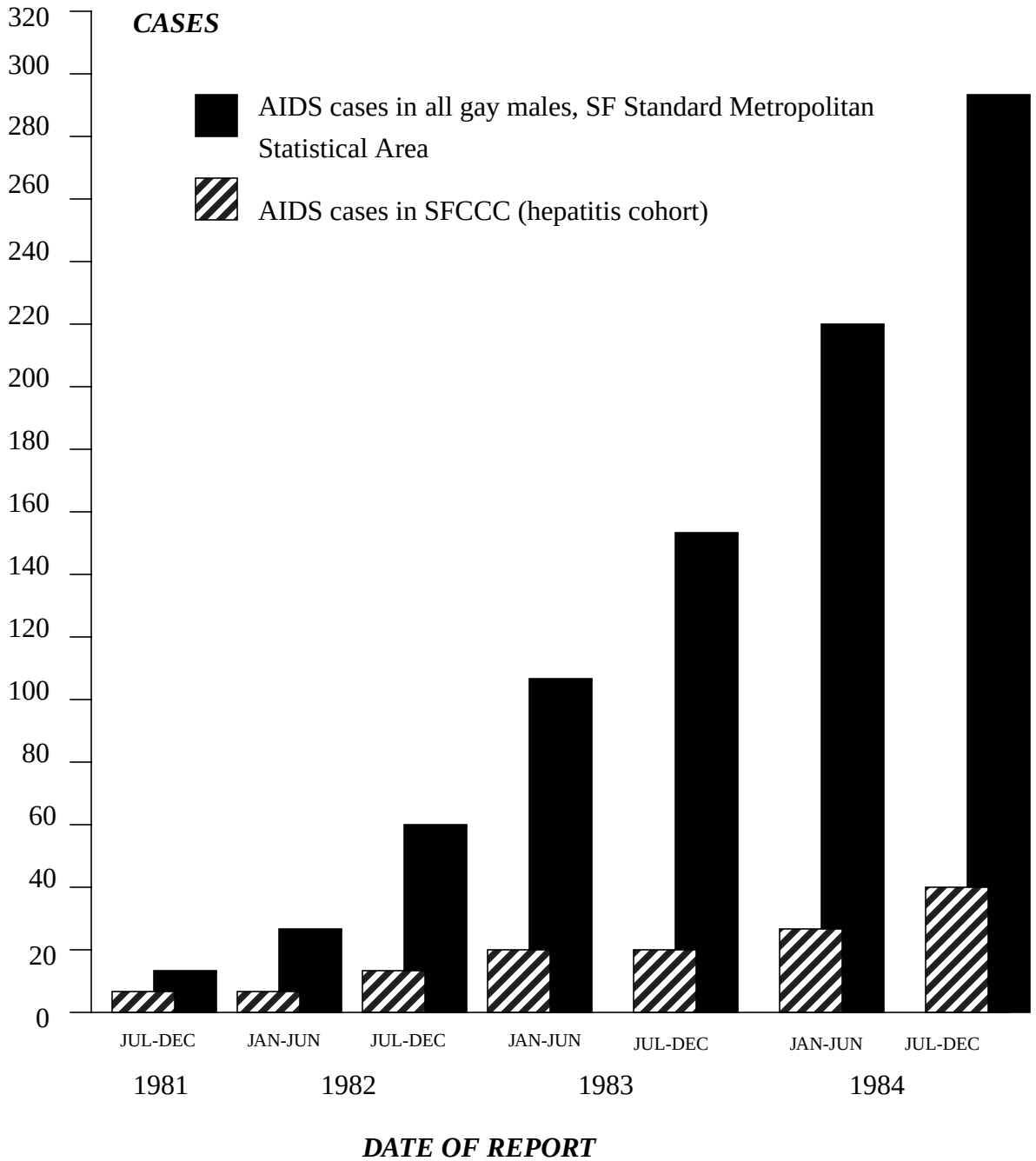

Figure 2

The preceding figure is taken from Jaffe, et al [10].

We would of course expect to find a larger absolute number of AIDS cases among gay men in the entire city than in the hepatitis trial, because there were approximately 8 times more gay males in the general population, than in the hepatitis trial. However, the ***growth rate*** was also consistently higher for the city in general, even in relative terms, during the early epidemic years. This fact is easily visible in the sharper curve, represented on the graph. This suggests that these first victims from the city as a whole were at even higher-risk for HIV infection, than the men in the SFCCC. Between 1981 and 1984, the numbers of new cases in the city increased more than 20-fold, while the numbers of new cases in the SFCCC increased less than 7-fold.

The ***overall*** gay population of the entire city should have been lower risk than the SFCCC, because the SFCCC was composed of men recruited at a VD clinic. However, the very first of the AIDS cases in the city do not merely reflect the overall population. They naturally tend to reflect a subgroup of the very highest-risk men in the entire city. This characterization is evidenced by the very fact of their becoming the first to be infected, out of the entire population. Men who have large numbers of partners are far more likely to be infected first. Exceptions to the rule would exist, but would be small in comparative numbers.

Thus, it is more than justified to treat the SFCCC and the total pool of men who contracted AIDS in these first few years, from 1981-1984 as being at least equal in risk for having acquired AIDS. This assumption is similar to that made previously for the Lemp study, but is even more conservative, because the “high-risk” population is defined using a much shorter period of time. The evidence of higher-risk in the general SF group is even more pronounced.

The fact that the men in the city as a whole had a sharper growth curve might at first seem to exonerate the hepatitis studies. However, this is contradicted by the significant over-representation of the SFCCC in the very first years. This over-representation suggests that the AIDS epidemic was somehow seeded initially among the SFCCC, and then spread like wildfire to the rest of the city.

There is clear evidence that the men from the overall gay population who became early AIDS cases were ***at least*** equal in risk for getting AIDS, in comparison to the men in the SFCCC. Thus, we can estimate statistically whether the higher representation of the SFCCC among the earliest AIDS cases could be attributed to random chance.

By the end of 1984, 166 of the SFCCC were diagnosed with AIDS, including 19 who had moved to other cities (Jaffe, et al [10]). The remaining group of 147 men is our “subgroup size”.

A total of 898 men in the San Francisco Standard Metropolitan Statistical Area (SMSA), which includes the SFCCC, were diagnosed with AIDS by the end of 1984. This represents our “total group size”.

We are interested in the sample of men who were the first cases, in 1981 thru 1982,

so we can compute the odds for the higher SFCCC representation to be the result of random chance.

The SFCC represented 60% of the first victims in 1981 [11] , then it quickly plummeted to 38.5% by the last half of 1981, and to 14.6% in the last half of 1984.

Six of the first 10 AIDS cases in San Francisco were members of the SFCCC [11], and 11 of the first 24 cases were also members [10]. By Jan 1, 1983, there were 104 AIDS cases in the city [12]. Of these, 35 were from the SFCCC [11] (estimated from bar graph).

- **6 of first 10 AIDS cases from SFCCC (1981, first half)**

Subgroup size = 147
Total group size = 898
Sample size = 10
n = 6
PROBABILITY IS: 0.00207447 (1 in 480)

- **11 of first 24 from AIDS cases SFCCC (1981, last half):**

Subgroup size = 147
Total group size = 898
Sample size = 24
n = 11
PROBABILITY IS: 0.00057291 (1 in 1700)

- **35 of first 104 AIDS cases from SFCCC (end of 1982):**

Subgroup size = 147
Total group size = 898
Sample size = 104
n = 35
PROBABILITY IS: 2.76209e-06 (1 in 362,000)

For 1981, the sample sizes are small, so it becomes more difficult to demonstrate highly dramatic differences (for very small samples, nearly ***any*** outcome tends to be possible, within reason). However, the probabilities are only in the range of a small fraction of a percent, so they are revealing enough, in themselves.

As the years progress and sample sizes increase, the differences become more evident, tending to confirm that the pattern is genuine and significant.

By the end of 1982, it is the most revealing, with a probability of less than 1 in 360,000 that the differences between the two groups were due simply to chance.

It should be noted also that 19 of the SFCCC men who developed AIDS are not reflected in the preceding statistics, because they had moved to other cities. If any appreciable number of these men were dropped from the charts before 1983, then it could significantly raise the levels of improbability, to a degree as high as 10 to the -15 (1 in 1,000,000,000,000,000)

It is important to note that men in the SFCCC were ***not*** methodically screened for AIDS in the early years of the epidemic. Thus, there was not a disproportionately higher level of reporting that would distort the comparison to other gay men. At the time of Jaffe's 6-year follow-up study [10], "No formal procedures were available to determine if patients who were reported to their health departments were cohort members." For the earliest AIDS cases, membership in the SFCCC was determined ***after*** the fact of diagnosis. These men had not received any experimental vaccine, so therefore had no special reason to be concerned about possible health effects. Epidemiological screening of the SFCCC for AIDS did not begin until 1983 [14].

Men in the SFCCC had the right to refrain from testing for AIDS, as some opted to do. None were forcibly tested.

4.2 High Rate of Seroconversion During Recruitment Years

This analysis hypothesizes that enrollment in the SFCCC seemed to spark a sudden onset of new HIV infection during the time-frame of 1978-1980, when men were recruited for this study.

Is this hypothesis born out by examination of the seroconversion dates for SFCCC participants? Estimated dates are provided by Rutherford, et al [14]. in the following, reproduced table:

Table 1, Estimated Year of HIV-1 Seroconversion in San Francisco City Clinic Cohort

Estimated year of seroconversion	New Seroconversions	
1977	3	
1978	**114**	
1979	**143**	**} Recruitment Years**
1980	**121**	
1981	42	
1982	30	
1983	7	
1984	9	
1985	4	
1986	5	
1987	3	
1988	4	
1989	4	

The new seroconversion figures represent a sample of 489 SFCCC men who had progressed to full-blown AIDS prior to the Rutherford study. These men were used to study the incubation period for AIDS, from initial HIV seroconversion.

The Rutherford study claims that 8% of a sample of 2877 men were HIV+ upon entry into the SFCCC study. Of the above 489 subgroup, 64% were supposedly seropositive upon entry, and 36% seroconverted within 24 months of entry.

The Rutherford study's estimate of the AIDS incubation period (about 11 years) is in line with current thinking. This suggests that the estimated dates of seroconversion are also accurate.

The fact that there were men who were claimed to be HIV-positive upon entry into the study does not exonerate the study's possible connection to the outbreak of AIDS.

Men were not recruited all at the same time for the SFCCC study. Most were recruited over a two year period, between 1978 and 1980. The first 831 participants were recruited over a 5 month period, between June and May of 1978 [14]. If men were infected continuously as they were recruited, they could also have been quickly spreading HIV through the general population at the same time, also infecting new recruits.

Even if you subtracted entirely from the graph all 312 men who were supposedly HIV+ upon entry, there would still be a noticeable bulge in the graph of seroconversions, around the years of recruitment. As computed from the table, this would still leave 69 seroconversions through the end of the recruitment period in 1980, another 72 in the next two year period of 81-82, and only 36 in the entire 7 year period following that, from 83-89.

This is particularly suspicious because rate of HIV spread should have continued to accelerate through much of the 1980s, because of the growing pool of already-infected men. It is well documented, how there were considerable difficulties in convincing the gay community to close the baths, for example, and to change sexual habits. Given the slowness and marginal results of this process, it is not reasonable to see such drastic fall-off in the rate of new infections in high-risk men, when the critical mass of already-infected men is now so much larger.

Even if many of these men were supposedly seropositive upon entry into the SFCCC, what would explain the extremely high clustering around the recruitment years or 1978-1980? This peculiarity requires explanation, in any case.

Not only is the sharp drop-off of new HIV seroconversions in high-risk gay men suspicious, but the paucity of seroconversions before 1978, when the trials started, is also peculiar. For example, there is no evidence that HIV infection existed in San Francisco before 1977. This includes no retroactive evidence of transfusion-related AIDS, in the context of totally unprotected blood banks, preceding 1977.

If there were accidental or intentional infection that had taken place in during the hepatitis study, stored blood samples could also have been contaminated at will, to obscure that fact.

Regardless of the necessarily-suspect claim of HIV upon entry into the cohort, the huge burst of seroconversion matching precisely the recruitment period is clear.

These statistical peculiarities suggest an unnatural, large-scale, simultaneous mass-infection that seeded HIV into the gay community of San Francisco.

4.3 Unreasonable Study Size/Selection

The government's criteria for choosing its hepatitis study participants and its sample sizes ought to seem questionable, in light of the basic rules of sampling.

The 6875 men in the SFCCC represent between 1 in 8 and 1 in 10 all the gay men in San Francisco, according to the Lemp estimates. Was such a large percentage of the gay population really necessary for this study?

If the government had chosen a sample of 1000 gay men to study, its margin of error in representing the total population would be 3.1% (95% confidence interval).

If they had chosen a sample of 3438 men (half the size of the SFCCC), the margin of error would be about 1.7% (95% confidence interval). By doubling the size of the sample, to 6875 men, reduces the margin of error only to 1.2%. This would represent a classic mistake in choosing sample size, by doubling the size and expense of your study, in return for only a slight extra return in accuracy.

This error would be compounded, however. One of the cardinal rules of sampling is that you must select a sample that is ***representative*** of the larger population that you wish to study. The assumption for your margin of error, as calculated above, depends completely on this. You cannot have "sampling bias", for example favoring young over old, rich over poor, urban over rural, etc, based on your method of picking study participants.

The SFCC was exclusively chosen from urban gay males, by way of a VD clinic. For monitoring purposes, the study would not have been representative of the general public, or of the general gay community, or even of the general gay urban male population. It included no heterosexuals, no women, no lesbians, no rural or small town gay males, and few urban gay males having more average sexual habits.

For "monitoring" purposes, what would be the point? So that it could be announced, with highest possible degree of accuracy, that in order to avoid hepatitis, you should avoid the obvious factors of having thousands of sex partners, and the extensive use of dangerous drugs? To announce that people who had already-known risk factors for hepatitis were indeed getting lots of hepatitis?

Furthermore, these hepatitis studies involved not only the full 6875 gay men in San Francisco, but also roughly 9000 gay men in New York City [13].

The most seemingly legitimate reason for the large screening program might be to identify a subgroup of high-risk men that had never been exposed to hepatitis B. Such men were needed as suitable subjects for the hepatitis vaccine trials.

However, even this does not quite make sense, because 3 to 5 times as many gay men were screened and found to be negative for hepatitis, than were actually used in the vaccine programs. In NY, about 10,000 men were screened, of which 3200 were HBV negative, of which only 1090 were enrolled for the vaccine tests. In San Francisco, 6704 were screened, of which 1676 were HBV negative, of which only 359 were enrolled for vaccine tests. Detailed questionnaires about sexual habits were issued ***after*** the vaccine selection was made [18]. It would have been possible to screen far fewer men in order to find the necessary vaccine trial participants.

It seems implausible that the government had an outpouring of concern for the health of promiscuous urban gay males, relating to a common disease (hepatitis) that involved comparatively few fatalities. When the fatal AIDS epidemic began spreading like wildfire in the gay community, the government appeared to be relatively unconcerned, for a long period of time.

It also does not appear likely that subjects were chosen for hepatitis studies merely to benefit those who were at highest risk for hepatitis. Alaskan Eskimos, including the Dena' Ina Tribe, were also chosen for hepatitis vaccine experiments [24]. They also alleged that they suffered health effects. Among their complaints was a charge that they were among the ***lowest*** risk for hepatitis, and were chosen merely for use as human guinea pigs, as an expendable population.

The selection of such a large and biased population for the hepatitis experiment was not consistent with sound rules of sampling, but would have been consistent with an intent to infect an "undesirable" population.

5 Epidemiological Anomalies

The following sections discuss various reasons why the early AIDS epidemic does not fit a reasonable epidemiological model.

5.1 Unreasonable Delay of HIV in IV Drug Community

According to the city of San Francisco [5], the total number of AIDS cases in the city, as compiled in 1999, showed about 79% gay males, 11% IV drug-using gay males, and 7% heterosexual IV drug users. Among the IV drug users, gay males were greater than half.

Because HIV is a blood-borne virus, crossover of HIV infections into the gay and straight IV drug community should have been rapid. Lemp estimates that 1120 gay males in the city were infected as early as 1979. Yet Lemp claims a 0% rate of HIV infection among drug users as late as 1981.

After even 100 HIV infections, the probability would be 99.9998% that at least one of those gay males would have been a drug user.

Furthermore, the infectivity rate of HIV due to a contaminated needle is nearly as high as the infectivity rate of unprotected anal sex [16][17]. Once that crossover occurred into the IV community, it should have spread quickly.

Lemp also estimated that 3% of the IV drug community would become infected, per year. Thus, the delay of at least 3 years (from 1978 to 1981) for crossover to have occurred in the IV drug users defies normal epidemiological expectations.

5.2 Anomalies Revealed by Computer Modeling

It might make intuitive sense to some observers that there are peculiar aspects to the manner in which the AIDS epidemic unfolded.

Throughout the 1970s there was no awareness of any such problem as AIDS, in spite of a supposed presence of HIV in human beings, as early as 1930. In the early 1980s, AIDS suddenly exploded with pronounced visible effect, almost simultaneously in far-flung locations round the globe: in Africa, in Haiti and the Caribbean, in Europe, in North America.

Retrospectively diagnosed AIDS cases from earlier in the 1970s exist at best anecdotally, in small handfuls. In some cases, even the anecdotes are subject to questions about reliability.

Epidemics cannot be modeled with great precision, because there are many variables that are complicated, or unknown, or unpredictable. However, computer modeling can determine if there

are profound inconsistencies in terms of parameters such as time, numbers of people infected, infectivity rates per sexual contact, numbers of sexual contacts, etc.

Appendix E gives the details of an analysis using epidemic modeling software that was developed by the author (a software engineer of more than 25 years experience). Included in this discussion are demonstrations that the software accurately duplicates real-life experimental results, as well as conforming with theoretical, mathematical testing.

The major conclusions of this analysis are as follows:

- The rate of HIV infection in the San Francisco Hepatitis B Vaccine Cohort is far higher than what could be reasonably expected
- Full-blown AIDS cases should have been evident many years earlier in San Francisco, based on the numbers of individuals infected in the early 1980s.
- In order to produce the levels of new infections seen per year in San Francisco, it would have required nearly 2000 infections to have existed as early as 1976- a time when virtually no HIV has been retrospectively discovered in the city.
- The numbers of men that would be required to appear suddenly in the late 1970s, in order to account for the subsequent level of HIV growth, is far more than can be reasonably accounted for by natural explanations such as vacationing in high-risk areas.
- Models attempting to explain the HIV growth curve by postulating variable rates of HIV infectivity are also inadequate.

The number of new HIV infections that can occur in a given year depends heavily on the size of the existing pool of already-infected persons. If you live in a city where only a handful of persons are infected, your odds of coming in contact with those few people is very small. If many thousands of people are infected, then your odds of encountering an infected person, and thus becoming infected yourself, are much greater.

This is why an epidemic “gathers steam” as it progresses, producing new cases per year at a faster pace, as more people become infected.

Similar to the adage that it “takes money to make money”, you could also say that it “takes infections to make infections”. The notion that small handfuls of infected people could spark a sudden explosion on the scale that was seen in the early 1980s, is profoundly inconsistent with existing knowledge about HIV infectivity rates. No matter how promiscuous those few individuals or their partners might have been, they could not produce thousands of new cases within a span of a few years, even if they had tried to do so.

6 Unreasonable Approval of the Vaccine

The commercially-made hepatitis B vaccine was considered as "safe, immunogenic, and efficacious" in a Sept. 1982 report by Dr. Don Francis of the CDC [15], without regard to the fact that gay men in the trials had started to become infected with an unknown, new disease, starting in 1981. Quick approval of the vaccine for general use was not prudent. Even if the vaccine had no connection to AIDS, the researchers would have had no way to be certain of that fact, until 1984, when HIV was discovered and could be detected in blood products.

Researchers were already well aware by the early 1970s of the existence of "slow" viruses [23].

For example, the visna virus, a sheep retrovirus with a long incubation period, had been discovered as early as 1949 (Fields, Virology, Chapter 55).

Scientists were also well aware of the dangers for vaccine contamination.

In an earlier vaccine fiasco, a potentially cancer-causing monkey virus (SV-40) had contaminated vaccines in the 1960s and was injected into millions of people [22].

Alaskan Native-Americans also claim to have been harmed by hepatitis vaccines[24]. The Yupik Eskimos and other Alaskan tribes were also used in hepatitis vaccine experiments. In a 1990 Council meeting, the Dena' Ina Tribe's Health Committee declared an "almost total loss of confidence" in the U.S. vaccine programs because of a wide variety of health problems, including AIDS-like symptoms, following hepatitis vaccinations.

In 1985, Dr. Don Francis also painted a rosy picture of the outcome of the hepatitis experiments on Alaskan natives, again calling them "safe, immunogenic, and efficacious" [25].

If the approval of the vaccine was not simply motivated by profit, at the expense of safety, then there is at least one other possible explanation. Perhaps some scientists were not concerned that test subjects were becoming infected, because they knew that the gay men had been infected intentionally with HIV during the experiment. This may be harsh speculation, but it is justified by the other wise inexplicable irresponsibility entailed in the premature approval of the vaccine, in the context of the circumstances and the state of knowledge that existed at that time.

7 Historical Context

The statistical analysis cannot be complete without reference to the political atmosphere in the late 1960s and the 1970s.

In a 1969 Congressional appropriations hearing for the Department of Defense, a Pentagon official named Donald MacArthur, a biological warfare expert, stated, “Within the next 5 or 10 years, it would probably be possible to make a new infective micro-organism which could differ in certain important aspects from any known disease-causing organisms. Most important of these is that it might be refractory to the immunological and therapeutic processes upon which we depend to maintain our relative freedom from infectious disease." [9] The proposed budget implied that this feat could be accomplished for the relatively modest sum of $10 million.

In the late 1960s, President Richard Nixon publicly renounced germ warfare, except for “defensive research.” In 1971 he ordered a large part of the army’s biological warfare unit at Fort Detrick, Maryland, transferred over to the nearby National Cancer Institute (NCI), where Dr. Robert Gallo would later discover the AIDS virus (HIV) in 1984. With the transfer of the biological war unit to the NCI, the army’s DNA and genetic engineering programs were coordinated into anti-cancer and molecular biology programs. It is quite possible that military biowarfare research could have continued under the guise of legitimate cancer research [9]

The Russians also signed the Biological and Toxin Weapons Convention in 1973, but immediately set up Biopreparat, a huge program for biowarfare research. Only in the late 1990s did the Yeltsin government admit to the existence of this secret program, which astonished American scientists with its scope, involving some 40 facilities. The Soviets embarked on this program in large part because they had believed that the United States had not ended its bioweapons program, but had simply hidden it away. A British intelligence officer recalled, “The notion that the Americans had given up their biological weapons program was thought of as the Great American Lie.” [26].

As far back as the 1950s, the United States maintained the ability to kill and incapacitate targeted people with biological weapons (see “In Search for the Manchurian Candidate”, John Marks [27]). The Technical Services Staff (TSS) of the CIA paid the Army Chemical Corp’s Special Operations $200,000 per year in return for operations systems to infect enemies with disease (Chapter 5, [27]). Dozens of germs and toxins were maintained for killing purposes.

Specific instances of assassination efforts using poisons and diseases are documented, targeting various figures including Fidel Castro and Patrice Lumumba of the Congo. A Newsday article reprinted in the Boston Globe(1/9/77) reports that CIA operatives received swine flu virus at a CIA biological warfare training station, and then attempted to spread the virus to Cuban pigs. Numerous other allegations of biological assaults against Cuban crops, livestock, and civilian populations are recorded.

Did the U.S. every really stop biowar activities, as Nixon claimed? The Congressional Church Committee hearings in the mid 1970s explored abuses in the CIA, and revealed that millions had continued to be spent on unauthorized biowar research.

Former CIA Director William Colby testified concerning a device called a “non-discernible microbioinoculator”, which was designed to deliver fatal injections of toxins, in such a way that could not easily be detected in an autopsy [27]. The CIA’s apparent goal was to develop various ways of killing that would leave no trace. It was also revealed that the CIA had stored enough shellfish toxin to kill a half-million people, an amount admitted to be far in excess of research needs.

Colby wrote in his memoirs that his admission about the “microbioinoculator” had “blown off the roof”, and led to his immediate dismissal. President Ford replaced Colby with future President George Bush, who gained a reputation as a staunch defender of CIA secrecy.

Undertones of violence existed in the Nixon administration, proven by his taped remarks urging the beating of anti-war protestors, and by G. Gordon Liddy’s admission in his autobiography of discussions contemplating the murder of political columnist Jack Anderson [28].

The Chicago Tribune published transcripts of Nixon’s Oval Office remarks about gays and other minorities. Nixon openly defamed Mexicans, blacks and Jews, saying Mexicans were prone to steal; blacks lived like “a bunch of dogs”; and homosexuals “destroyed” strong societies. Nixon spoke of the historical need for the Catholic Church to “clean out” its homosexuals. He admired societies that tried to eliminate homosexuals, saying, "Let’s look at the strong societies. The Russians. Goddamn, they root ‘em out. They don’t let ‘em around at all.” Deploring the alleged takeover of San Francisco by "fags", Nixon proclaimed, “...I can’t shake hands with anybody from San Francisco.”

The climate of hatred against gays intensified in the 1970s with thousands of homosexuals coming out of the closet with unprecedented political demands. It is easy to imagine that they might have been vulnerable to covert government-sponsored medical experiments, similar to those secret radiation experiments that had been conducted on unsuspecting citizens during the Cold War years, up to the year 1974, when the government’s investigation of the records documenting these crimes ended.

During the 1970s, there was extensive animal retrovirus experimentation undertaken as part of Nixon’s “War on Cancer”. Animal viruses were transferred between species and manipulated genetically. As a result, new cancerous and immunosuppressive diseases were produced experimentally. Could the AIDS virus have arisen from this dangerous and unprecedented experimentation? Is it possible that anecdotal cases of pre-1960s AIDS in Africa were misdiagnosed, or might have represented false positives, or contaminated blood samples? Could such cases have been contrived as a “cover-up” in order to discredit research pointing to a man-made origin of the AIDS epidemic, in vaccine programs of the 1960s and early 1970s in Africa?

In the 1970s Don Francis and Max Essex (later to become top scientists in AIDS) experimented extensively with feline leukemia virus (FELV), an HIV-like retrovirus that produced a disease in cats, similar to human AIDS. Early in the AIDS epidemic, both scientists suspected that AIDS might be caused by a retrovirus, because the new disease was so reminiscent of these earlier studies. [30] .

No human retrovirus was known until 1978 when Robert Gallo discovered a retrovirus that caused a rare type of human leukemia. In the early 1980s, a second leukemia virus was discovered by Gallo. These human "T-cell" leukemia viruses were termed HTLV-1 and HTLV-2. When Gallo discovered HIV in 1984, he initially called it "human T-cell leukemia/lymphoma virus", or HTLV-3. Later, the term "human immunodeficiency virus". (HIV) was substituted,

which represented only a change in name.

After Don Francis completed his work with the HIV-like cat retrovirus, he joined the Centers for Disease Control and headed the hepatitis B vaccine experiments, using gay men as guinea pigs in San Francisco and other cities (the very same experiments discussed throughout this document) [31].

This document has demonstrated a statistical correlation between the government-sponsored hepatitis experiments and the outbreak of HIV in the gay male volunteers. The historical context reinforces the concern that the correlation might have resulted from intentional infection of these men. There is proof of extreme bigotry in high office. There is a proven record of intrigue, deception, political corruption, unauthorized experimentation, and use of human guinea pigs. There was a great deal of relevant scientific research which suggests that it would have been possible for our government to have either discovered or created HIV, before the hepatitis trials began. This combination of circumstances makes the requirement to reinvestigate the hepatitis experiments all the more compelling.

8 Conclusions

In the early 1980s, gay men who were given experimental hepatitis B vaccine showed significantly higher rates of HIV infection, compared to other gay men of equally high risk for HIV infection, in the SFMHS group. The odds for these differences being due to random chance alone are extremely small, by some measures in a range of one-millionth of one-millionth.

The justification for saying that the gay men in the SFMHS group were of equally high risk is based not merely on their characterization as such, in the Lemp study, or merely on the fact that the men were chosen from high-risk tracts within San Francisco. It is based on comparison of HIV growth patterns within the two groups, which were highly similar for many years in the study period. It is also based on the fact that the SFMHS group "caught up" and even surpassed the vaccine group in HIV infection, in spite of the "head start" that the vaccine group had in the early years. This suggests that the non-vaccine SFMHS group may have been even ***higher*** risk. Therefore, the vaccine group should by rights have shown noticeably ***lower*** prevalence of HIV, much less the higher levels that they actually showed, compared to the SFMHS.

The SFCCC group showed even stronger evidence of being lower risk in comparison to the "control" group to which they were compared, which consisted of other early AIDS victims in the city. The higher initial rate of AIDS in the SFCCC showed that mere participation in the hepatitis studies, in any capacity, had a statistically significant association with AIDS diagnosis, beyond what could be explained by risk alone.

In the past, the high rate of HIV/AIDS among men in the hepatitis study has been attributed to their "high risk" status, or to random chance, or else it has been denied that the rate of HIV/AIDS was in fact higher than that of the general population. This analysis has shown those rather simplistic explanations to be inadequate.

With these significant possibilities effectively ruled out, it is still conceivable that one might try to find alternative, benign explanations. However, the search for benign explanations might begin to strain the imagination, no less than the thought of contaminated vaccines, or intentional infection by some covert, unknown means.

Because the production of the vaccine involved use of pooled blood from high-risk gay men, the possibility of HIV-contaminated vaccines has sometimes been imagined as a tragic accident.

However, vaccine (Heptavax-B produced by Merck Sharp & Dohme) was inactivated using three steps: pepsin, urea, and formaldehyde (formalin) (Francis et al., 1986) (see reference [6]):

> ***"In this study, we demonstrate that each of the three inactivation steps used in the manufacture of Heptavax-B independently will inactivate the infectivity of high-titered preparations of the AIDS virus"***

If this claim is correct, then not even the use of pooled blood from gay men should have caused vaccine contamination.

Furthermore, samples of the vaccines were tested retroactively for HIV. It was claimed that no HIV was detected in the vaccines.

An important question is why there appears to be unusually high rates of HIV/AIDS associated with mere participation in the hepatitis studies, even for men who received no vaccine. These differences clearly do not seem attributable merely to a higher risk status.

The hepatitis trials in New York City would be worthy of similar analysis. In 1980, the rates of HIV infection among vaccine trial participants were in the range of 20%, even higher than the rate in San Francisco [9].

All possible explanations deserve consideration and investigation, even the most politically sensitive explanations. Evidence suggests that accidental infection should have been unlikely, yet a nearly undeniable statistical correlation remains to the vaccines and the hepatitis studies. To the degree that accident can be ruled out, the possibility for intentional infection is strengthened.

Significant numbers of people do not approve of homosexuals, and would not object to their removal from society. Historically, “undesirable” populations, such as prisoners, mentally retarded and others were used for unethical experimentation. Even if no such criminal malice existed in the hepatitis studies, the best way to establish that fact would be to treat the possibility seriously, and investigate it thoroughly.

If HIV is an old virus, present long before the 1970s, then this would only make it easier to suppose that the virus could have been discovered without public announcement, and then tested on an undesirable population.

In the 1990s, horrific details of the government’s Cold War experimentation during the 1940s up to the 1970s came to light. Thousands of covert radiation experiments were performed on children, the mentally ill, hospitalized patients, pregnant women, Native Americans, and other U.S. citizens. Thus, during the 1970s, it would not have been unprecedented for government scientists to experiment on gay men, the most hated minority in America.

Our government should long ago have made full disclosure as to the fates of the men who volunteered for these experiments. Exactly how many men who received the experimental vaccines died of AIDS, and in what years?

Why has the scientific community failed to notice these profound statistical correlations? It is understandable that these realizations might have escaped the notice of laymen, but they were well within the ability of trained scientists to discern easily. The fact that they did not, even after allegations of a vaccine link, is evidence that these allegations have not been taken seriously enough.

It is essential for the scientific community to explain the high rates of HIV/AIDS in the hepatitis study members.

There is no suggestion being made here that starting a man-made epidemic would have been anything other than an act of madness. It would be madness unprecedented in scale, but not historically unprecedented in its recklessness or cruelty. It was little short of madness how the Reagan administration essentially ignored the new disease, tried to slash the CDC budget, and ordered the Surgeon General not to mention the word "AIDS" in public. If a government was capable of ignoring a new disease to such a degree, it is only a short step further to infer that they might have been foolish enough to precipitate the disease. It may have been an act of irrational religious fervor. Perhaps it was imagined, or foreseen, that the virus would be confined largely to risk groups. Perhaps the perpetrators simply did not care about ***any*** citizens who did not live according to strict Christian sexual morality. Perhaps the perpetrators, even if Americans, were infiltrated by foreign enemies who might have wished the entire country's destruction.

It is not a necessity for this study to explain the exact nature of the madness, but merely to document why an act of madness has very probably occurred.

The connection between the origin of the AIDS epidemic and the government experiments has been dismissed by the AIDS epidemic. However, the statistical analysis presented here demonstrates a definite correlation. A reopening of the entire matter is in order.

The purpose of this document is not to cast final judgement concerning the origin of AIDS in the gay community. It is to demonstrate that there is a strong and suspicious link, which has no obvious explanation, between the outbreak of AIDS and the government-sponsored hepatitis studies. The purpose is to call for investigation of this important question.

9 Refuting Counter-Arguments

This section will review and refute various attempted criticisms concerning the link of AIDS and the vaccine studies.

- **Perhaps the men in the vaccine trial had sex with each other, and infected each other**

  It is true that the men signing up for a trial might be friends who also have sex with each other. However, this factor could be equally true for the SFMHS (non-vaccine group), as for the SFHBVCS (vaccine group).

  The sponsors of a vaccine trial would not publish lists of everyone involved in the trial. Fraternizing among participants would likely be limited to small groups of friends who might have known each other prior to the trial. It would not likely be intermingling among the entire cohorts.

  The men were also chosen because of a high-risk profile, meaning that they were regular clients of institutions such as bars and baths. They would have been prone to sexual encounters with large number of the general gay male population, not merely with a small set of friends.

- **Perhaps the men in the vaccine group became complacent because of feeling "protected" by the vaccine, and started practicing more risky behavior**

  The analysis already demonstrated that risk behavior was if anything, even higher in the "control" (non-vaccine) group that was used for comparison.

  The vaccine would have protected only against hepatitis, even ***if*** it worked. It would not have protected against a host of other venereal diseases, including herpes, syphilis, and gonorrhea. The sponsors of the trial would have been remiss if they had not explained this to the men involved.

- **AIDS has a 10-year incubation period. Therefore, the men must have been infected prior to the start of the trial, because most became infected in less than 10 years.**

  The Lemp study was specifically measuring the dates of initial HIV infection, and ***not*** the development of full-blown AIDS. The incubation period is not relevant, in this context.

  The fact that we are looking at HIV seroconversion rather than "AIDS" also eliminates another argument used in the past against efforts to suggest a vaccine connection: that a reaction to the vaccine challenged the immune system, and hastened the development of AIDS. Here, we are looking only at virus exposure, not disease symptom development.

  In the case of the SFCCC, where the analysis was examining initial rates of AIDS diagnosis,

estimates of the HIV seroconversion dates were provided by the Rutherford study. The dates were heavily clustered around the dates of recruitment into the hepatitis study.

- **Perhaps the men receiving the vaccine were simply monitored more closely, and their HIV status was detected more quickly than for men who did not receive the vaccine**

For the Lemp study, this line of argument clearly does not apply. Both groups, vaccine and non-vaccine, had stored blood samples that were taken annually and later examined for HIV.

- **Perhaps the men in the vaccine group were actually higher risk than the non-vaccine group**

There are degrees of "high-risk". Simply because the non-vaccine group was characterized in the Lemp study as "high-risk", it does not necessarily mean that they were equally as high risk as the men in the vaccine group. However, this analysis is not relying merely on subjective descriptions. It assesses their risk level by comparing their actual patterns of HIV growth. The non-vaccine control group, the SMHS appeared, if anything, to be even ***higher*** risk for HIV infection.

Furthermore, the SFHBVCS vaccine group was chosen only from men who had no previous exposure to HBV (hepatitis B virus), and were chosen from men who seemed in good health at the outset of the trial (else, it would have invalidated the vaccine trial results). HBV was sexually spread and epidemic among gay men, which was part of the reason that gay men were chosen for the vaccine trial, in the first place. One would expect that there would be a significant correlation between HBV exposure and HIV exposure. The exact extent of this factor is difficult to measure, but it is nonetheless an additional, qualitative reason to believe that the men in the vaccine trial might have been actually lower in overall risk for ***previous*** HIV exposure, at the start of the vaccine trial.

Dr. Lemp's study itself states, "Although these vaccine cohort members were recruited from sexually transmitted disease clinics, they represent lower-risk since none of the cohort members were seropositive for hepatitis B virus at time of recruitment." Dr. Paul M O'Malley, Project Director of the SF Dept of Health AIDS research study, also concurred, "Their blood had not in 1980 shown signs of infection with hepatitis B, which can be spread through sexual activity. The subjects were therefore assumed to be less sexually active than other SF clinic visitors." If it is necessarily true that the men were less active, perhaps it would be more accurate to say that they had at least beaten the odds, in terms of encountering infected partners.

- **HIV growth is an exponential function. This can magnify differences over time between groups, compared to what you would see with a flat rate of infection, such as by exposure to carcinogens.**

It has been demonstrated that the hepatitis/vaccine study groups had, if anything, even less

risk for HIV than the "control" groups used for this comparative analysis. All of the men in both groups are from the same geographical area, visiting the same limited number of bars/ baths in the city, showing similar behaviors. It is therefore justified to treat the hepatitis/vaccine study groups as being ***at least*** equal in risk, with any error in that assessment being weighted against a conclusion that implicates a correlation to the hepatitis study.

As long as the characterization of equal risk levels is accurate, then the earliest AIDS cases, which represent an essentially random sample taken from the total pool of men, should not disproportionately reflect either of the subgroups.

■ **Perhaps the statistical sample size is too small in order to draw conclusions**

This objection is not applicable to the probability calculations of "n or more" in this document. Similarly. if a coin is flipped with 30 heads in a row, it is not "too small" of a number of trials in order to draw a confident conclusion. This point is discussed further in the general statistical primer, Appendix F.

Where "sample size" becomes an issue is in the question of whether the SFMHS group is sufficiently large in order to use as an estimate of the HIV rate for all high-risk men.
The Lemp study in fact uses SFMHS for the whole gay male population of the city, not distinguishing between "higher/lower" risk. The study states, "Since the SFMHS is a population-based probability sample, its seroprevalence estimates are likely to be representative of HIV seroprevalence for homosexual and bisexual men in San Francisco."

As a double-check, we can take the year of 1982, and do another calculation based on a reversed hypothesis: suppose that the 40% rate of HIV infection in the SFHBVCS group, rather than the 23% HIV infection rate of the SFMHS group, was the "real" rate of HIV for high-risk men. What would the probability then be that the 799 men in the SFMHS might show their 23% rate of HIV infection, by random chance alone? Might this be within the bounds of normal possibility?

In this case, our subgroup size is 799 and the total group size is still 35000. The sample size (the number of men that we expect to have HIV in 1982) is 40% of 35000 =14000 . Of these 14000, the SFMHS group represents (23% x 799) = 184 men. What we want is the probability that 184 ***or fewer*** would be in the SFMHS group. Our program calculates "or greater", so we can simply calculate the odds for "185 or greater", and subtract this from one.

Using the program, this produces an answer of zero, or essentially no chance (too small of a value to represent). Thus, it does not matter which end of the spectrum that we choose as representing HIV prevalence in the overall gay population. The difference between the vaccine and control groups in either case is more than what we could expect by random chance.

■ **Perhaps the data concerning HIV infection rates are not reliable**

HIV antibody tests are more than 99% accurate [7]. The false-positive rate for the Elisa HIV antibody test is only 1 to 5 per 100,000 assays.

The false negative rate is only 1 in 450,000 to 1 in 660,000 [8] . It can be assured that these tests must be reliable, because they have had long use in protecting our nation's blood supply.

These error rates have no significant impact on the computed probabilities.

Furthermore, any false-positives or false-negatives would have tended to affect both the vaccine and non-vaccine cohorts in ***equal*** proportions, thereby tending to cancel out, in any case.

In the absence of better data for the period prior to 1982, it is justified to use the Lemp data as representing the best of what is currently available. The conclusion derived from this data would also represent the best conclusion that could be drawn at this time.

10 About the Author

I (Tom Keske) am a gay/AIDS activist in the Boston area, originally from Ohio. I have been a data communications software engineer for 25 years, B.S.(with honors) in Computer Engineering, Case Western Reserve University, ‘74. Education includes modest background in statistics, strong in math and programming. I am an internet activist, writing frequently about AIDS and its origins. I am age 48, and in a committed relationship with a lifetime partner of 28 years, Daniel.

Academic honors include National Merit Scholar and National Honor Society.

I was “survey statistician / computer programmer” responsible for telecommunication software at the Bureau of Census, in Washington, D.C., from 1974 - 1979, where I received an “outstanding employee” award. I was President of the Board of Directors, Bradbury Park Condominiums, in the Maryland suburbs of Washington. In 1979, I relocated to the Boston area, where I am currently employed as a senior staff engineer for a Fortune 500 vendor of communications equipment. Between 1980 and the present, I have worked on development of statistical multiplexors, intelligent matrix switches, a data PABX, Ethernet bridges, multiprotocol routers and cable routers.

I have also worked in the area of data encryption, and have performed statistical analysis of encryption methods. In the late 1990’s, I successfully defended a self-designed encryption program in the face of a $1000 reward posted to break the code. During college, I had faced discrimination at the National Security Agency during a job interview at Fort Meade, with polygraph questions about homosexuality.

I am a supporter of progressive causes, such as in past participation as an Amnesty International “Freedom Writer”.

My activities instilled a keen awareness of political corruption and abuse, which led to an interest in research and criticism of the CIA and intelligence establishment.

As a gay activist, I engaged in public vigil/hunger strike in front of the Massachusetts State House in support of the gay civil rights bill, which passed immediately afterward, after more than 15 years of effort. There was modest coverage in smaller, local papers. A state legislator said that the protest was lending moral weight to the cause. I have also engaged in civil disobedience protests, outside the Supreme Court, during the first March on Washington, and in various cities including Raleigh, North Carolina and Atlanta, Georgia.

I am health-conscious, and for hobbies enjoy swimming, hiking, bicycling, and chess.

Thomas R. Keske, 205 Warren St., Randolph, Mass. 02368

email: tkeske@mediaone.net (781) 961-1571

11 References

[1] ***The River: a Journey to the Source of HIV and AIDS***, by Edward Hooper (former BBC Africa correspondent) (Penguin; Little, Brown, New York, 1999)

[2] HIV RESEARCH SECTION, SAN FRANCISCO DEPARTMENT OF PUBLIC HEALTH, http://www.dph.sf.ca.us/php/hivresearch.htm

[3] Lemp GF, Payne SF, Rutherford GW, Hessol NA, et al:
Projections of AIDS morbidity and mortality in San Francisco,
JAMA 1990 Mar 16;263(11):1497-501 PMID: 2407871; UI: 90172481

The abstract of the Lemp 1990 study reads as follows*:*

Abstract: To develop a model for predicting acquired immunodeficiency syndrome (AIDS) morbidity in San Francisco, Calif, through June 1993, we combined annual human immunodeficiency virus seroconversion rates for homosexual and bisexual men and for heterosexual intravenous drug users with estimates of the cumulative proportion of the population with AIDS by duration of human immunodeficiency virus infection and with estimates of the size of the at-risk populations. We projected AIDS mortality by applying Kaplan-Meier estimates of survival time following diagnosis to the projected number of cases. The median incubation period for AIDS among homosexual and bisexual men infected with the human immunodeficiency virus was estimated to be 11.0 years (mean, 11.8 years; 95% confidence interval, 10.6 to 13.0 years). The model projects 12,349 to 17,022 cumulative cases of AIDS in San Francisco through June 1993, with 9,966 to 12,767 cumulative deaths.

[4] University of San Francisco (home page, includes links to faculty):
http://www.ucop.edu/srphome/uarp/

[5] San Francisco HIV/AIDS Statistics as of Nov. 28, 1999,
http://www.sfaf.org/aboutaids/statistics/index.html

[6] Francis DP, Feorino PM, McDougal S, et al. *The safety of the hepatitis B vaccine. Inactivation of the AIDS virus during routine vaccine manufacture,* JAMA 1986 Aug 15;256(7):869-72.

[7] Centers of Disease Control NAC, from Guide to Information and Resources on HIV Testing, 1997

[8] CDC FAQ: http://www.cdc.gov/nchstp/hiv_aids/pubs/faq/faq15.htm

[9] ***AIDS and the Doctors of Death: An Inquiry into the Origin of the AIDS Epidemic*** (1988), and ***Queer Blood: The Secret AIDS Genocide Plot*** (1990), by Dr. Alan Cantwell, Jr., Aries Rising Press, PO Box 29532, Los Angeles, Calif. 90029 (AlanRCan@aol.com)

[10] Jaffe HW, Darrow WW, Echenberg DF, et al.: *The Acquired Immunodeficiency Syndrome in a Cohort of Homosexual Men, A Six-Year Follow-up Study,* Annals of Internal Medicine. 1985;103:210-214

[11]Centers For Disease Control, Morbidity and Mortality Weekly Report, Sept. 27, 1985 / Vol 34 / No. 38

[12]Moss AR, Bacchetti P, Osmond D et al: *Incidence of the Acquired Immunodeficiency Syndrome in San Francisco,* 1980-1983, Journal of Infectious Diseases, Vol 152, No. 1, July 1985

[13]Koblin BA, Morrison JM, Taylor PE, et al.: *Mortality Trends in a Cohort of Homosexual Men in New York City, 1978-1988,* American Journal of Epidemiology, Vol 136, No. 6

[14]Rutherford GW, Lifson AR, Hessol NA et al: *Course of HIV-1 infection in cohort study of homosexual and bisexual men: an 11-year follow up study,* Br Med, Vol 301, Nov 24, 1990

[15]Francis DP, Hadler SC, Thomppson SE, et al: *The prevention of hepatitis B with vaccine. Report of the Centers for Disease Control multi-center efficacy trial among homosexual men* Ann Intern Med 1982 Sep;97(3):362 PMID: 6810736, UI: 82282328

[16] Kaplan EH, Heimer R, *A model-based estimate of HIV infectivity via needle sharing,* J Acquir Immune Defic Syndr 1992;5(11):1116-8, Yale, PMID: 1403641, UI: 93020182

[17] Vittinghoff E, Douglas J, Judson F, et al: *Per-contact risk of human immunodeficiency virus transmission between male sexual partners,* Am J Epidemiol 1999 Aug 1;150(3):306-11, PMID: 10430236, UI: 99357305

[18] Padian NS, Shiboski SC, Glass SO, Vittinghoff E:
Heterosexual transmission of human immunodeficiency virus
(HIV) in northern California: results from a ten-year study,
Am J Epidemiol 1997 Aug 15;146(4):350-7 PMID: 9270414, UI: 97416464

[19]Godfried JP, Hessol NA, Koblin BA, et al: *Epidemiology of Human Immunodeficiency Virus Type 1, Infection among Homosexual Men Participating in Hepatitis B Vaccine Trials in Amsterdam, New York City, and San Francisco, 1978 - 1990,* Amer Journal of Epidemiology, Vol 137, No .8, 1993.

[20] University of Southern California, http://hivinsite.ucsf.edu/akb/1997/01txbld/index.html#Ba

[21] Jacques JA, Koopman JS, Simon CP, Longini IM: *Role of primary infection in epidemics of HIV infection in gay cohorts*, J Acquir Immune Defic Syndr 1994 Nov, PMID: 7932084, UI: 95017548

[22] "The Polio Vaccine and Simian Virus 40", by By T.J. Moriarty, http://www.chronicillnet.org/online/bensweet.html

[23] "Retroviruses- An Introduction", JAMA HIV/AIDS Information Center, http://www.ama-assn.org/special/hiv/newsline/briefing/retro.htm

[24] "Alaska Health Issues and Indigenous Peoples" (video), Mary Ann Mills, Bernadine Atchison, Delice Calcote, July 1991 Arctic Village Health Conference. These women are Activists against medical experimentation on Alaska Native communities.

[25] Heyward WL, Bender TR, Francis, DP et al: *The control of hepatitis B virus infection with vaccine in Yupik Eskimos Demonstration of safety, immunogenicity, and efficacy under field conditions*, Am J Epidemiol 1985 PMID: 3160233, UI: 85248405

[26] "The Bioweaponeers", the New Yorker, March 1998, by Richard Preston

[27] ***In Search of the Manchurian Candidate***, by John Marks, 1988, Times Books, ISBN: 0-440-20137-3 Senator Edward Kennedy said of this expose,"John Marks has accomplished what two U.S. Senate committees could not".

[28] ***Secret Agenda***, by Jim Hougan, Random House, 1984, ISBN: 0-394-51428-9. The Los Angeles Times called this book "a monument of research and fact-finding". Hougan was Washington Editor of Harper's magazine, and helped produce the Emmy Award winning documentary, "Confessions of a Dangerous Man"

[29]"The Role of Robert Gallo in the Origin of AIDS", Kwame Ingemar Ljungqvist, http://homepage.calypso.net/~ci-15476/toa/gallo.html. Mr. Ljungqvist is editor the Swedish scientific journal, "Science of the 21st Century"

[30] 11/96 "1 in 10 Talk Show" interview with Max Essex

[31]Harvard Public Health Review, "The Gathering Storm", by Sarah Abrams http://www.hsph.harvard.edu/review/the_gathering.shtml (describes the career of researcher Don Francis)

12 Acknowledgments

Much thanks to Dr. Alan Cantwell, Jr., author of “Queer Blood (1993)” and “AIDS and the Doctors of Death” (1988) [9] for help in editing this document and reviewing the facts. Dr. Cantwell has spent some 14 years investigating the hepatitis vaccine experiments.

Thanks also to Billi Goldberg, San Francisco AIDS researcher/activist, whose discussions of the 1990 Lemp study inspired this further statistical analysis.

13 Document Reproduction

Appendix A Demonstrating the Validity of the Statistical Approach

The logical approach used in this document is to compare groups of gay men who received vaccines, or who otherwise participated in hepatitis studies, with other groups of similar gay men who did not engage in these activities. This type of statistical analysis is virtually identical to how researchers compare their own test vaccine group to a control/placebo group.

Researchers are typically trying to prove that their vaccine group showed statistically lower incidence of the targeted disease, compared to the control group. Or, perhaps, they might try to show that the vaccine group showed statistically higher levels of protective antibody response, compared to the control group.

The only difference in this document's use of the same technique is that the aim is to show the statistical ***presence*** of another disease, instead of the ***absence*** of the disease that the vaccine tries to prevent. The "control" groups are defined retrospectively, by identifying groups of men who were of demonstrably equal or higher risk for HIV/AIDS.

The following is an abstract of a Swiss study involving hepatitis B vaccine. This will demonstrate how the vaccine researchers are using very similar statistical calculations:

> *Evaluation of tolerability and antibody response after recombinant human granulocyte-macrophage colony-stimulating factor (rhGM-CSF) and a single dose of recombinant hepatitis B vaccine.*
>
> *Tarr PE, Lin R, Mueller EA, Kovarik JM, Guillaume M, Jones TC*
>
> *Sandoz Pharma Ltd, Basel, Switzerland.*
>
> *Recombinant human granulocyte-macrophage colony stimulating factor (rhGM-CSF) has been shown to augment antigen presentation by macrophages and dendritic cells in vitro, and to increase antibody responses to injected antigens in experimental animals. To evaluate the usefulness of rhGM- CSF as a vaccine adjuvant, 108 healthy volunteers were randomly assigned to receive an injection of rhGM-CSF (n = 81) or placebo (control group; n = 27), followed by an injection with recombinant hepatitis B vaccine into the same site. During the study period of 28 days, protective antibody titers to hepatitis surface antigen (anti-HBs10 mIU ml-1) were observed in 11 of 81 subjects receiving rhGM- CSF, but in none of the controls (P = 0.035). Injections were well tolerated. A single i.m. or s.c. injection of 20- 40 micrograms of rhGM- CSF significantly enhances antibody responses when given at the same site as recombinant hepatitis B vaccination.*
>
> *Publication Types: Clinical trial Randomized controlled trial*

PMID: 8961505, UI: 97120835

The study involved 81 people in a vaccine group, and 27 in a placebo (control) group, for a total of 108. In the vaccine group, 11 showed protective antibodies, but none did in the placebo group. When the researchers say (P = 0.035), they mean that the probability of this outcome is 3.5%.

Using the same program in [27], we get the same result:

Subgroup size = 81
Total group size = 108
Sample size = 11
n = 11
PROBABILITY IS: 0.035144

Note that the researchers must make assumptions that are virtually identical to what we must make in evaluating whether the vaccine was causing HIV infection. When they try to prove that the vaccine prevented hepatitis B infection, they must assure that the vaccine group is at equal risk for hepatitis, compared to their placebo group. There cannot be differences in age, general health, risk of exposure, etc, that might account for the different outcomes between the two groups. The researchers must compute statistically that the differences in hepatitis rates between the two groups are not merely a result of "random chance".

There is no such thing two absolutely identical groups, but the process of "averaging out" can mitigate the effects of small differences.

The Francis, et al study evaluated the gay hepatitis B vaccine using the same type of analysis [15]. It estimated the effect of the vaccine based on about 907 men in a vaccine group , and 495 in a placebo group (this is including other cities as well, not just San Francisco). It had to assume that these men were roughly equal in risk for acquiring hepatitis, just as we have had to demonstrate equal risk of men for acquiring HIV.

The Francis study concluded that the vaccine was beneficial in preventing hepatitis B on the strength of 56 new hepatitis infections in the control group, versus only 11 in the vaccine group (probability = .0004, or 1 in 25000). These results are not nearly as compelling as the figures cited in this document linking HIV infection to the hepatitis studies and vaccines.

Yet, the Francis figures were used to justify dispensing the vaccine to millions of people, whose health and life would be in the balance. It is therefore difficult to argue that the figures in this document are not on as solid of a scientific basis, and adequate justification to draw a conclusion.

Appendix B Letter From Dr. George Lemp

>Dear Mr. Keske:

Thank you for your interest in my research. The data cited appear reasonably accurate. I assume the author looked at the published graphs and guessed the approximate data points. My time doesn't allow me to try to dig up the original data points, but I took another look at the graph and the author's guesses seem reasonable (perhaps off by 1% at a few points). Who was the author and where were these data cited? The JAMA article would be on file at any University or major hospital medical library in your area. JAMA is widely held by libraries and should be available. If you have trouble finding it, please email your address and we'll mail you a reprint.
Sincerely,

George Lemp

X-Sender: uarp@popserv.ucop.edu
>X-Mailer: QUALCOMM Windows Eudora Pro Version 4.2.0.58
>Date: Mon, 31 Jan 2000 09:38:08 -0800
>To: george.lemp@ucop.edu
>From: Universitywide AIDS Research Program <uarp@ucop.edu>
>Subject: Fwd: 1990 Study Data
>
>
>>From: "Thomas Keske" <TKeske@mediaone.net>
>>To: <uarp@ucop.edu>
>>Subject: 1990 Study Data
>>Date: Fri, 28 Jan 2000 23:24:48 -0500
>>X-Mailer: Microsoft Outlook Express 5.00.2919.6600
>>
>> Jan. 28, 2000
>>Dr. George Lemp
>>University of California
>>uarp@ucop.edu
>>

>>Dear Dr. Lemp,
>>
>>I hope that my emailing will not impose on your time. I much
>>appreciate all the work that you have done for AIDS research.
>>I have a very brief question, and would much appreciate if you
>>could reply.
>>
>>I am trying to follow a thread on sci.med.aids, which quoted data
>>from your 1990 study, showing rates of HIV in the early 1980's:
>>
>> SFHBVCS: 1978 - 0.3%, 1979 - 4%, 1980 - 15%, 1981- 28%, 1982 - 40%,
>> 1983 - 46%, 1984 - 47%, 1985 - 48%, 1986 - 48%, 1987 - 49.3%
>>
>> SFMHS: 1978 - 0%, 1979 - 2%, 1980 - 4%, 1981- 10%, 1982 - 23%,
>> 1983 - 42%, 1984 - 48%, 1985 - 49%, 1986 - 49.3%, 1987 - 49.3%
>>
>> SFHBVCS = San Francisco City Clinic Cohort Study
>> SFMHS = San Francisco Men's Health Study
>>
>>The author said that this data was extracted from charts in the 1990 study,
>>but I have been unable to locate the full text of the study:
>>
>> Lemp GF, Payne SF, Rutherford GW, Hessol NA, Winkelstein W Jr, Wiley JA,
>> Moss AR, Chaisson RE, Chen RT, Feigal DW Jr, Thomas PA, Werdegar D.
>> Projections of AIDS morbidity and mortality in San Francisco. JAMA 1990
>> Mar 16;263(11):1497-1501
>>
>>Could you please tell me if the data above appears to be
>>reasonably accurate, or how I could obtain the full study?
>>I am asking only as an interested layman.
>>
>>Thanks very much for your time.
>>
>>Regards, Tom Keske
>>

Appendix C Error Analysis for Lemp Data Calculations

C.1 Effect of Variation in High Risk Population Estimate

It might seem at first glance that the calculation could be in error if the estimate of the size of the "high risk" gay population is too high. In actuality, it turns out that lowering the estimate of the total high risk population size will work to ***decrease*** the probability that the result could be attributed to random chance. The is because reducing the estimate of the high risk population size also lowers the "sample size" of HIV+ men that we expect to draw in any one year (23 percent of the total high-risk men, for the example year of 1982).

Below is a listing of the computed probabilities for different values of the estimated number of high-risk gay men in San Francisco, ranging from as many as 100000, to as few as 10000.

As it can be seen, these variations matter little in the resulting probability:

```
Subgroup size = 359, Total group size = 100000, Sample size = 23000, n = 144
PROBABILITY IS: 3.3273e-13

Subgroup size = 359, Total group size = 50000, Sample size = 11500, n = 144
PROBABILITY IS: 2.98706e-13

Subgroup size = 359, Total group size = 30000. Sample size = 6900, n = 144
PROBABILITY IS: 2.58323e-13

Subgroup size = 359. Total group size = 25000, Sample size = 5750, n = 144
PROBABILITY IS: 2.40079e-13

Subgroup size = 359, Total group size = 20000, Sample size = 4600, n = 144
PROBABILITY IS: 2.14932e-13

Subgroup size = 359, Total group size = 15000, Sample size = 3450, n = 144
PROBABILITY IS: 1.78359e-13

Subgroup size = 359, Total group size = 10000, Sample size = 2300, n = 144
PROBABILITY IS: 1.21827e-13
```

The ***largest*** of these probabilities is roughly 1 in 3,000,000,000,000.

C.2 Effect of Errors in HIV Infection Rate Figures

There could have been errors in reading the chart data from the Lemp study. Dr. Lemp had suggested that this might have amounted to a percent or so (Appendix B).

The following computations test the effect of errors in the probability calculation for 1982, by reducing the number of HIV+ men in the vaccine group, and increasing the number of HIV+ in the non-vaccine group, in increments of 1%. Both of these adjustments work to increase the probability that the outcome might be attributable to random chance. The computations range from 1% adjustments, to 3% adjustments. This was only an informally suggested error rate, so it is being tripled for safety, with worst case assumed jointly for each affected variable:

- **Adjusting sample size +1% and vaccine group size -1%**

```
non-vaccine = 8400    (24% of 35000)
vaccine     =  140    (39% of 359)

Subgroup size = 359, Total group size = 35000, Sample size = 8400, n = 140
PROBABILITY IS: 1.63957e-10
```

- **Adjusting sample size +2% and vaccine group size -2%**

```
non-vaccine = 8750    (25% of 35000)
vaccine     =  136    (38% of 359)

Subgroup size = 359, Total group size = 35000, Sample size = 8750, n = 136
PROBABILITY IS: 4.00199e-08
```

- **Adjusting sample size +3% and vaccine group size -3%**

```
non-vaccine =  9100   (26% of 35000)
vaccine     =   133   (37% of 359)

Subgroup size = 359, Total group size = 35000, Sample size = 9100, n = 133
PROBABILITY IS: 2.41677e-06
```

Allowing for errors of +3% in the HIV+ sample size and -3% in the number of HIV+ vaccine group men (jointly) has the effect of improving the odds that the correlation could be a product of random chance, but not to a significant degree (worst case of roughly 1 in 400,000)

Appendix D Letter from Case Western Reserve Statistics Department

The following note from the CWRU Statistics Dept was in response to a query that I made (appended), trying to validate the general reasoning behind the analysis. In this query I had rephrased the question, to avoid biasing, as one of evaluating the safety of a "food additive" (instead of a vaccine) that was being tested as a possible "carcinogen" (cancer-causing, instead of AIDS-causing). I posed the question using the identical numbers from the vaccine analysis, for the year 1980 in the Lemp study, for group size, sample, size, etc. Below is the email exchange:

From: Joe Sedransk
Department of Statistics, CWRU
Cleveland, OH 44106-7054

Mr. Keske: Your reasoning is mostly correct. The main assumption is that in the absence of the food additive the mortality rate would be 4%; that is, that the lab animals are "similar" to the general population (with a mortality rate of 4%). [I'd ask how the mortality rate of 4% was determined.] Then you would find the probability of 54 or more deaths out of the 359 lab animals, assuming a mortality rate for each of 4%. (There is also an assumption that the events (life/death) are independent among the 359 animals. This would usually be true, but should be verified.) I did a crude calculation using a normal distribution approximation (approximating what you did) and found that the probability of 54 or more deaths is extraordinarily small. The only problem with your formulation is that it is 54 or more out of 359 rather than out of 1400. I hope that this helps.
Sincerely,
Joe Sedransk

At 11:29 PM 2/4/00 -0500, you wrote:
> Jan. 31, 2000
>Dear Mr. Sedransk,

> I am an alumnus of the CWRU class of '74, in Computer
> Engineering.

> I was wondering if it would be too much trouble if you could help
> to clarify my understanding of a simple type of statistics problem,

> or if you could direct me to another resource. I am trying to
> understand how to evaluate the following type of
>problem:

> A group of 359 lab animals using a food additive showed
> a 15% rate of cancer in a year (=54 animals). The normal rate
> of cancer, measured in a total population of 35000 such animals,
> was 4% (=1400 animals)
>
> QUESTION: Is this a normal statistical variation, or should
> it be judged that the food additive is unsafe?

> It seems to me that this is similar to a problem where you
> have 35000 marbles in a bag, 359 are black and the rest white.
> If you draw a random sample of 1400 marbles, what is the probability
> of getting 54 or more black marbles, by random chance alone?
> This can be computed from the binomial distribution curve. I've computed
> the probability as 5.6 x (10 to the -17). Therefore, my conclusion is
> that the food additive should almost certainly be suspected as carcinogenic, and should
> not be approved for mass consumption. This question came up only as
> part of a newsgroup discussion (nothing related to business and school).
> We were (embarrassingly) unable to agree on the answer.
> Could you please help us to settle it?
>
> Regards,
>Tom Keske (class of '74)

NOTE: the issue of how the 4% figure was derived is explained earlier in the document (basically, taking the HIV rate of the high-risk SFMHS group of gay men as being representative for other high-risk gay men in San Francisco).

The 54 men of the 359 in the SFHBVCS (vaccine) group of gay men is the number that actually acquired HIV by 1980 (15% of the group, per the Lemp study, as opposed to the roughly expected value of 4%). The "random sample" of all HIV positive men for 1980 is 1400 (4% of 3500). In this 1400 is included the 54 from the group of 359 SFHBVCS men. Later, I chose 1982 rather than 1980 as the main focus, because the data was measured rather than extrapolated.

Appendix E Software Epidemic Modeling Analysis

Various epidemiological anomalies concerning the origin and spread of HIV can be demonstrated through the use of computer modeling software.

The modeling software referenced in this section was developed by the author. The sources are not listed here because of length (more than 1500 lines, for two programs), but are available on request from the author. The programs can run on a PC with Microsoft Visual C++, or any ANSI standard C compiler.

The software is general-purpose and flexible, capable of using any modeling assumptions that the user might to make, and allowing various different models to be tested.

The programs do not contain built-in assumptions about parameters such as infectivity rates or degrees of risk behavior. These parameters are defined by means of an interactive dialogue, when the program is run.

The vepid.c software is capable of specifying any number of risk subgroups that engage in particular mixes of activities, at different frequencies. Infectivity rates may be time-varying, to account for factors such as the stages of HIV infection, where the first few weeks might involve higher infectiousness.

The program also ask for initial rates of HIV prevalence, when the modeling period begins.

The program keeps track of each individual member of the modeled population, and whether they are currently infected, or not. The members of the population are randomly paired for sexual contacts, according to their risk group's quota for the year, evenly spread throughout the months of each year. Partners are chosen based on their willingness for a compatible activity.

When an uninfected person is paired with an infected person, the program decides whether the uninfected person will become infected or not. This is random, but is kept strictly within the bounds of the average probability for infection, based that person's role in the current contact.

The program prints totals of new and cumulative infections, for each year.

The epid.c program is a simpler and faster epidemic modeling package that models a single population group, with up to two specified active/passive activities.

The modeling examples that follow will all take a number of measures to produce "worst case" projections for HIV growth, when looking at the general gay population:

- Both receptive and insertive sexual roles are treated as having the same infectivity as the more risky "receptive" role.
- No account is made for monogamous or partially monogamous partners. All members of the target population are treated as if being promiscuous.
- No account is made for safe practices, such mutual masturbation, dildos, condoms, etc.
- The rate of sexual activity is assumed for the entire population is assumed to be as high as for the "high risk" men in the vaccine trials.

E.1 Per-Contact Infection Rates

Published figures exist for per-contact probabilities of infection for various sexual acts/roles with infected partners[16][17][18]. This makes it possible to do computer modeling to examine the spread of HIV. Average rates of infectivity, per-contact with an HIV+ partner, are:

- Anal receptive: .0082
- Anal insertive: .0067
- Oral receptive: .0006
- Oral insertive: 0 (slight, theoretical only)
- Vaginal, male-to-female: .0009
- Vaginal, female-to-male: .0001125

In order to model, you also need data as to the sexual practices and frequencies in the target populations. For the hepatitis B vaccine trial participants in San Francisco, this is listed as 67 contacts with different partners per year [18]. For New York gay men in the vaccine trials, the rate was lower, at about 40 partners per year. Virtually all men reported a mix of both anal and oral sex, with lower-risk oral sex being somewhat more prevalent.

Since these figures are for high-risk men, it would be a generous over-estimate to apply the same rate to all gay men in the city.

E.2 Evidence of Program Accuracy

E.2.1 Consistency with Independent Mathematical Test

As a check whether the program is working correctly, we can try a simple case that approximates a coin-flipping problem, where we can compute the expected answer by another means. Say that you have a population of 100000 where half are infected. Say that the probability of "infection" is 50% (= .5), and that these people pair up for a single sex act (50000 pairings). How many should be infected?

The program says:

```
% epid
Enter Population Size ( <= 100000): 100000
Enter number initially infected: 50000
Probability, infection per ACTIVE  contact, type #1: .5
Probability, infection per PASSIVE contact, type #1: .5
Probability, infection per ACTIVE  contact, type #2: .5
Probability, infection per PASSIVE contact, type #2: .5
Enter average number of contacts per year: 1
Enter number of years: 1
Enter random seed (any number between 1 and 4294967295): 987987347
New infections in year #1 = 12442, GRAND TOTAL = 62442
```

You might suppose that the expected value of new infections is (50000 * .5) = 2500. However, you must take into account that the pairings are random, not simply pairings of infected persons with uninfected persons. When an infected person is paired with another infected person, or an uninfected person is paired with another uninfected person, nothing changes. The question is, how many pairings of uninfected and infected persons should there be? The answer that we really expect is about half of the number of pairings of infected + uninfected partners.

This is computed using the "combinations" function, described earlier. The total pairings are

(100000 C 2) = 5e+09. Pairings of two infected or two uninfected partners would each amount to ((50000 C 2) / 5e+09) = 25%. The mixed pairings would constitute the remaining 50%.

Thus, we should expect roughly (50000 * 0.5 * 0.5) = 12500 new infections, versus the program's projected 12442, which is very close to expected (within bounds of expected, random variation).

E.2.2 Consistency with Real-Life Experimental Results

A California study of heterosexuals [18] followed 360 heterosexual woman who were HIV negative, but had regular male partners who were HIV+. These women continued to have unprotected sex with their male partners. In a ten-year period, the study reported 68 new HIV infections among the 360 women.

The vepid.c modeling software comes reasonably close in attempting to duplicate the results of the California study. Starting with 360 infected men, the program reported 73 new infections among the women- very close to the reported value of 68.

Below is the output of the "vepid.c" epidemic modeling software, for this experiment.

Subgroup #1 represents the 360 males (all initially infected). Subgroup #2 represents the infected men's 360 female partners (0 initial infections).

The abstract of the study did not list a number of sexual contacts per year, so I made a conservative estimate of one intercourse every other week (26 contacts per year).

```
% vepid
Enter Total Population Size ( <= 100000): 720
Enter no. of activities to model:  1
DO ANY INFECTIVITY RATES VARY WITH TIME (y or n)? n
Does activity #1 involve exactly 2 partners (y or n)? y

Enter av probability of infection,
       activity #1, ACTIVE  role: .0001125
Enter av probability of infection,
       activity #1, PASSIVE role: .0009

DO YOU WISH TO DEFINE POPULATION RISK SUBGROUPS (y or n)? y
ENTER NUMBER OF POPULATION RISK SUBGROUPS: 2
Enter size of subgroup #1: 360

Enter activity #1, ACTIVE , average no. contacts per year
       for subgroup #1:  26
Enter activity #1, PASSIVE, average no. contacts per year
       for subgroup #1:  0

Enter number initially infected for subgroup #1:  360
Enter size of subgroup #2: 360

Enter activity #1, ACTIVE , average no. contacts per year
       for subgroup #2:  0
Enter activity #1, PASSIVE, average no. contacts per year
       for subgroup #2:  26

Enter number initially infected for subgroup #2:  0
```

```
Enter number of years to model: 10
Enter random seed (any number between 1 and 4294967295): 987987987

NUM ACTIVITIES: 1
ACTIVITY #1, ONE_PARTNER
      AV prob infection, ACTIVE , 0.0001125    num_adjust,  = 0

      AV prob infection, PASSIVE, 0.0009       num_adjust,  = 0

POP SIZE: 720
NUM SUBGROUPS = 2

   TOTAL FOR SUBGROUP 0 = 360
   TOTAL INFECTED IN SUBGROUP: 360

   Contacts/yr for activity #1, ACTIVE : 26

   TOTAL FOR SUBGROUP 1 = 360
   TOTAL INFECTED IN SUBGROUP: 0

   Contacts/yr for activity #1, PASSIVE: 26

New infections in year #1 = 13, GRAND TOTAL = 373
New infections in year #2 = 9, GRAND TOTAL = 382
New infections in year #3 = 8, GRAND TOTAL = 390
New infections in year #4 = 9, GRAND TOTAL = 399
New infections in year #5 = 7, GRAND TOTAL = 406
New infections in year #6 = 6, GRAND TOTAL = 412
New infections in year #7 = 6, GRAND TOTAL = 418
New infections in year #8 = 3, GRAND TOTAL = 421
New infections in year #9 = 7, GRAND TOTAL = 428
New infections in year #10 = 5, GRAND TOTAL = 433
TOTAL CONTACTS: 187200
TOTAL DUMMY CONTACTS, NO PARTNER: 0
REDUNDANT INFECTIONS: 12
Subgroup #1 infections: initial = 360, new = 0, total = 360
Subgroup #2 infections: initial = 0, new = 73, total = 73
```

E.3 First Year, SFHBVCS

The first modeling is of the 359 men in the vaccine trial, who went from .3% infection (1 man) in 1978 to 4% infection (14 men) in a single year. Is this rate suspiciously high? The following treats the 360 men as a “closed” population, having sex with each other (which should make cases rise even faster). The average number of partners was increased from 67 to 104, to make it even more conservative. The program output follows:

```
% epid
Enter Population Size ( <= 100000): 360
Enter number initially infected: 1
Probability, infection per ACTIVE  contact, type #1: .0082
Probability, infection per PASSIVE contact, type #1: .0082
Probability, infection per ACTIVE  contact, type #2: .0006
Probability, infection per PASSIVE contact, type #2: .0006
Enter average number of contacts per year: 104
Enter number of years: 1
Enter random seed (any number between 1 and 4294967295): 24525
New infections in year #1 = 2, GRAND TOTAL = 3
```

The observed number of infections was nearly 5 times higher than expected by a generous modeling estimate.

To double check that the program's modeling is not simply too low, we can try another set of years, with a larger initial pool of infected men. In 1982, 40% of the 359 were infected (144 men). By 1983, 46% were infected (165 men). The percent of the total population becoming newly infected is higher (6% versus 3.7%) and the absolute numbers of men newly infected is higher (21 versus 14). The only difference is the pool of men initially infected.

For this, the program shows:

```
% epid
Enter Population Size ( <= 100000): 360
Enter number initially infected: 144
Probability, infection per ACTIVE  contact, type #1: .0082
Probability, infection per PASSIVE contact, type #1: .0082
Probability, infection per ACTIVE  contact, type #2: .0006
Probability, infection per PASSIVE contact, type #2: .0006
Enter average number of contacts per year: 104
Enter number of years: 1
Enter random seed (any number between 1 and 4294967295): 24525
New infections in year #1 = 42, GRAND TOTAL = 186
```

The program in this case shows ***more*** men being infected that actually observed, demonstrating that it is not simply a matter of the infectivity/frequency estimates that causes our previous low value. What makes the difference is the number *initially* infected.

The computer model is saying that in order to have extremely high rates of new HIV growth, it is necessary to have a significantly large initial pool of infected men. High rates of HIV growth are not feasible in a scenario where only a small handful of men are initially infected. When such an unreasonably high rate of HIV growth is observed, it suggests that some mechanism exists to spread the virus that is above and beyond simply a high rate of sexual contact.

E.4 Patient Zero Scenario

For this test, the program estimated the course of HIV growth over a 20 year period, starting with a single, infected person (a "Patient Zero" type of scenario), for a gay population of 100,000, having behaviors similar to high-risk San Francisco men.

```
% epid
Enter Population Size ( <= 100000): 100000
Enter number initially infected: 1
Probability, infection per ACTIVE  contact, type #1: .0082
Probability, infection per PASSIVE contact, type #1: .0082
Probability, infection per ACTIVE  contact, type #2: .0006
Probability, infection per PASSIVE contact, type #2: .0006
Enter average number of contacts per year: 67
Enter number of years: 20
Enter random seed (any number between 1 and 4294967295): 4536356356
New infections in year #1 = 2, GRAND TOTAL = 3
New infections in year #2 = 5, GRAND TOTAL = 8
New infections in year #3 = 4, GRAND TOTAL = 12
New infections in year #4 = 5, GRAND TOTAL = 17
New infections in year #5 = 4, GRAND TOTAL = 21
New infections in year #6 = 7, GRAND TOTAL = 28
New infections in year #7 = 12, GRAND TOTAL = 40
New infections in year #8 = 10, GRAND TOTAL = 50
New infections in year #9 = 12, GRAND TOTAL = 62
New infections in year #10 = 19, GRAND TOTAL = 81
New infections in year #11 = 37, GRAND TOTAL = 118
New infections in year #12 = 39, GRAND TOTAL = 157
New infections in year #13 = 55, GRAND TOTAL = 212
New infections in year #14 = 64, GRAND TOTAL = 276
New infections in year #15 = 81, GRAND TOTAL = 357
New infections in year #16 = 144, GRAND TOTAL = 501
New infections in year #17 = 168, GRAND TOTAL = 669
New infections in year #18 = 242, GRAND TOTAL = 911
New infections in year #19 = 351, GRAND TOTAL = 1262
New infections in year #20 = 433, GRAND TOTAL = 1695
```

By the end of the 4th year of the 20-year period, there would have been about 50 infections. In 10 more years, most of these cases would have progressed to full-blown AIDS. At that time, there were still only about 250-300 infections, total.

When AIDS broke out, it took only a few dozen usual cases of Kaposi's Sarcoma before it was apparent to the medical establishment that there was an unusual problem. Thus, a realistic model of HIV growth says that the AIDS epidemic should have become apparent, at a time when HIV prevalence was still quite low. The rapid saturation of HIV in the gay community, subsequent to

the initial outbreak of AIDS, points to the fact that there was a mass, simultaneous infection of a larger number of men.

E.5 Estimated Seed Size in SF

How many initial, simultaneous, mass infections would have to suddenly appear in the late 1970s, in order to account for the rates of explosive growth that followed? For this discussion, this is what is meant by the "seed size".

Lemp's data shows near zero infection in the San Francisco gay population in 1978, rising to 49.3% for the entire gay male population of the city, by 1987. This would be approximately (56000 * .493) = 26708 HIV infections in 9 years.

The first reported case of transfusion AIDS in San Francisco was in 1982, 4 years after the start of the hepatitis study recruitment. Blood supply screening began in 1985. The first retroactively estimated case of transfusion related HIV infection was 7 years earlier, in 1978, also coinciding with the start of the hepatitis study [20].

To be generous, we can push the date of essentially-zero HIV prevalence to 1976, the year before the first back-dated projections of gay HIV seroconversions listed for high-risk men, cited by Rutherford [14].

Approximately how many men would it take for a seed size, in order to get 26708 HIV infections by 1987 (11 years)?

This can be estimated by running the modeling program repeatedly, taking an initial guess, and then working up or down, in iterative attempts.

As it turns out, the necessary seed size in 1976, as a conservative estimate, would need to be between 1900 and 2000 men:

```
% epid
Enter Population Size ( <= 100000): 56000
Enter number initially infected: 1900
Probability, infection per ACTIVE  contact, type #1: .0082
Probability, infection per PASSIVE contact, type #1: .0082
Probability, infection per ACTIVE  contact, type #2: .0006
Probability, infection per PASSIVE contact, type #2: .0006
Enter average number of contacts per year: 67
Enter number of years: 11
Enter random seed (any number between 1 and 4294967295): 11324234
New infections in year #1 = 625, GRAND TOTAL = 2525
```

New infections in year #2 = 845, GRAND TOTAL = 3370
New infections in year #3 = 1070, GRAND TOTAL = 4440
New infections in year #4 = 1351, GRAND TOTAL = 5791
New infections in year #5 = 1732, GRAND TOTAL = 7523
New infections in year #6 = 2097, GRAND TOTAL = 9620
New infections in year #7 = 2616, GRAND TOTAL = 12236
New infections in year #8 = 3046, GRAND TOTAL = 15282
New infections in year #9 = 3487, GRAND TOTAL = 18769
New infections in year #10 = 3906, GRAND TOTAL = 22675
New infections in year #11 = 3949, GRAND TOTAL = 26624

E.6 From Where Comes the Seed?

There is no recorded evidence of extensive HIV infection in the gay community, anywhere in America, in the mid-1970s. Traveling and vacation within the U.S. borders could not be a sufficient factor to account for simultaneous mass infection of 2000 men in San Francisco, merely in the space of a few years.

Immigration from other U.S. cities in time period also could not explain the number of men simultaneously mass-infected, given the lack of evidence for any appreciable degree of HIV elsewhere in the country.

The fact of a few anecdotal cases of supposed HIV infection from earlier years, such as a case claimed in 1959 in St. Louis, do not alter this fact. Having a few stray cases, even if these are not simply myths, does not create a scenario to allow rapid infection of a large number of men within a few years.

The same is true even of travel and vacation to other foreign locations, such as Africa or Haiti. As an example, following is an estimate of what would happen if nearly the entire gay population of San Francisco vacationed in Haiti for a couple weeks, mingling with 100% infected men, and having an average of 3 sexual contacts during that vacation:

% epid
Enter Population Size (<= 100000): 100000
Enter number initially infected: 50000
Probability, infection per ACTIVE contact, type #1: .0082
Probability, infection per PASSIVE contact, type #1: .0082
Probability, infection per ACTIVE contact, type #2: .0006
Probability, infection per PASSIVE contact, type #2: .0006
Enter average number of contacts per year: 3
Enter number of years: 1
Enter random seed (any number between 1 and 4294967295): 222345
New infections in year #1 = 352, GRAND TOTAL = 50352

In other words, a grand total of only 350+ infections. Of course, nowhere near the whole gay male population of SF men is going to vacation in Haiti in the space of a couple years, nor will this number immigrate.

It is a "Catch-22" which forbids large, sudden simultaneous mass infections: it requires a large pool of existing infections. To make a large pool of infections takes time, when you are starting from only a few infections. In this necessary time, AIDS would reveal itself much earlier. The degree of apparent seeding shows an artificial nature, more consistent with a hypothesis of unnaturally produced mass infection, such as in the hepatitis experiments.

E.7 Variable Infectivity Per Stage

A 1994 study at University of Michigan (Jacquez, et al) [21] attempted to explain the sharp rise and rapid fall-off of HIV infection, stating that "Thousandfold differences in transmission probabilities by stage of infection are needed to fit the epidemic curves". Their hypothesis was that the initial infection stage, characterized by flu-like symptoms, would cause an infectivity rate 1000 to 3000 times higher than the rate of infectivity in the "long, asymptomatic phase", lasting 10 years or more.

The per-contact risk for anal sex is broken out in this study as "0.1-0.3 per anal intercourse in the period of initial infection , 10(-4) to 10(-3) in the long asymptomatic period, and 10(-3) to 10(-2) in the period leading to AIDS."

When fed into the program, allowing a 1-month duration for an initial infection period, and using the higher of the Jacquez figures in each case (= 0.3 for initial infection, .001 for asymptomatic) this still did not appear to explain the pattern of HIV growth in the early epidemic years. In fact, when starting with a single infection, over a span of 11 years, the variable-rate infectivity figures actually came out to be lower than flat rate figures used in the previous examples (trying 56000 men, 200 initially infected, for 11 years). Only in the first year did the Jacquez figures produce a higher rate of infections (113 versus 59). In later years, it fell off sharply (after 5 years, 699 total for variable infectivity, 795 for flat rate; after 11 years, 1988 for variable rate and 3992 for flat rate).

This is probably because the high rate of .1-.3 only applies to an initial infectivity period that is very short (2-8 weeks). After that, the listed infectivity figures for Jacquez are actually lower (.001 versus .0082), for a much longer period of time (10 years).

There are additional, possible objections to the notion that a high rate of infectivity in the "initial infection" stage could account for the early explosion of AIDS. This initial period is

characterized by symptoms such as headache, vomiting and diarrhea This is not a scenario where even promiscuous gay men are likely to seek, or succeed in finding, a lot of partners.

If the body were that overcome with huge amounts of virus, the symptoms might be more severe than merely flu-like (one might imagine that the person would be dying). The authors of the Michigan study acknowledged that the questions of viral load during the different stages were a matter of controversy.

If the body is that vulnerable to massive proliferation of virus, when initially exposed, then it is more difficult to understand why the asymptomatic phase, with lower viral load, should be all that less infectious. It would seem only to require a small amount of virus to cause infection, if the virus can duplicate that quickly and freely in an unprepared host.

In the case of IV drug injection, the amount of virus is far less than in a typical amount of semen, yet the infectivity is very efficient. This also suggests that the amount of virus required for infection would not necessarily need to be great.

It is possible also that the infectivity of HIV has changed over time. It is not in the best evolutionary interest of a virus to kill its only natural host. It is a commonplace phenomenon for viruses to becomes less virulent over time. However, if such drastic changes have occurred merely within the last 20 years, then it might suggest that HIV was a relatively “new” virus, and detract from the likelihood that it has been infecting humans since the 1950s or 1930s.

More studies concerning viral loads at different stages of infection, and concerning the effect of viral load on infectivity of unprotected sex, would be useful.

There is a risk of circular reasoning in the example of the University of Michigan study- a “good” model must fit the observed curve. Thus, the model cannot test whether the observed curve is a natural phenomenon- that is an implicit assumption .

At this point, it appears more likely that researchers are stretching to find explanations for the high initial rates of HIV spread, and the sharp drop-off that followed. Perhaps part of the reason for these contortions is the refusal to examine a more controversial hypothesis: an artificial, mass seeding of HIV infections into the gay population.

Appendix F General Statistical Primer

There is sometimes a naive tendency to assume the cynical attitude that "you can prove anything with statistics." However, you do not live in a world of truths that possess absolute certainties. You live in a world that is probabilistic in nature. For better or worse, statistics is one of your best tools for assessing truth and falsehood in the world around you. Your ability to be deceived by statistics, or to be enlightened by it, will depend in large part on how much effort you put into acquiring a good command of the subject, so that you can critically evaluate statistical claims.

The vaccine analysis is similar to a problem in drawing samples of marbles of different colors from a jar. Suppose that you had a jar with 100 white marbles and 100 black marbles, evenly mixed. You close your eyes, and draw a sample of 20 marbles at random.

Because there are equal numbers of white and black marbles in the jar, you would expect, on average, that the 20 that you draw would also be roughly, evenly mixed- about 10 white marbles and 10 black. Of course, this is only an average, that you would expect to find over many similar trials. In any one sample of 20 marbles, you might have a few more white, or a few more black. This is a normal variation by random chance, or what you would call "luck of the draw".

Take an extremely simple case- a jar has one white marble and one black marble. You close your eyes and draw one marble. What is the chance that it will be white?

When this problem is fed into the computer program listed in the appendix, the answer is:

Chance of drawing a white marble, when pulling one marble at random from a jar containing 1 white marble and 1 black marble:

Subgroup size = 1
Total group size = 2
Sample size = 1
n = 1
PROBABILITY IS: 0.5

A probability of .5 means 50%, or an even 50-50 chance, the same as the chance of getting "heads" when flipping a coin.

When you flip a coin, or draw from two colored marbles, there is no particular reason that one outcome would be favored over another, so you tend over many trials to get half (50%) of each result. This same logic applies to ***any*** two events where there exist no logical reasons to favor one outcome over the other. If you randomly painted an "X" on half of the people in a room, and "Y" on the other half, then blindly chose a random sample of people in the room, you would expect to get about half "X" and half "Y" on average, regardless of how the initial selection had happened to have been made.

The same is true of the "random" grouping of men who received a vaccine, versus those who didn't. It should be simply irrelevant as a factor when drawing samples of men "chosen" at random to become HIV infected, as if you had merely painted "X" or "Y" on them.

If you have a factor that ***does*** favor an outcome, such as having a coin weighted on one side, then it is no longer a 50-50 chance. You have to be very careful not to have such hidden bias.

You also have the opportunity, if you see someone flipping a coin and getting 30 heads in a row, to know that something is fishy. It is probably a trick coin, because the probability would be only about 1 in a billion, by random chance. It is similar with our vaccines- something is clearly fishy. What exactly it might be, we have to investigate, but there is something that needs investigation.

The problem of drawing marbles from a jar is a bit different than flipping a coin. If you keep flipping a coin, the odds of getting "heads" is the same on every flip. If we pull a second marble from our jar, we are sure to get the white, because it is the only marble left. Our program should show this: if we draw two marbles, we should have 100% percent chance of getting a white marble:

Chance of drawing a white marble, when pulling two marbles from a jar containing 1 white marble and 1 black marble:

Subgroup size = 1
Total group size = 2
Sample size = 2
n = 1
PROBABILITY IS: 1

This is a case so obvious as to be nearly silly, but it demonstrates how marble-drawing problems (hypergeometric distributions) sometimes differ significantly from coin-flipping problems (binomial distributions). It also helps to confirm that our probability-computing program is working correctly. The program comes in handy, because the computations become extremely lengthy, when the mixtures of marbles and the tested conditions become more complicated.

What would be the odds of finding 10 or more white marbles, when drawing 20 marbles from a jar holding 100 white marbles and 100 black marbles? This is the most-expected result.

Chance of 10 or more white marbles, drawn in a sample of 20 marbles, randomly pulled from a jar of 100 white marbles and 100 black marbles:

Subgroup size = 100
Total group size = 200
Sample size = 20
n = 10

PROBABILITY IS: 0.592851

That is, there is about a .59, or a 59% chance of getting 10 or more white marbles.

Since this is the most expected result, we should see a slightly smaller probability for drawing 11 or more white marbles, and even less for drawing 12 or more white marbles. The computer program output below shows that this is true:

Chance of 11 or more white marbles, drawn in a sample of 20 marbles, randomly pulled from a jar of 100 white marbles and 100 black marbles:

Subgroup size = 100
Total group size = 200
Sample size = 20
n = 11
PROBABILITY IS: 0.407149

Chance of 12 or more white marbles, drawn in a sample of 20 marbles, randomly pulled from a jar of 100 white marbles and 100 black marbles:

Subgroup size = 100
Total group size = 200
Sample size = 20
n = 12
PROBABILITY IS: 0.240184

We had a 59% chance of drawing 10 or more white marbles. To draw 11 or more white marbles, the chances fall to 40%. To draw 12 or more white marbles, the chances fall to 24%.

What are the chances that the entire group of 20 marbles will be 100% all-white?

Chance of finding 20 all-white marbles, drawn in a sample of 20 marbles, randomly pulled from a jar of 100 white marbles and 100 black marbles:

Subgroup size = 100
Total group size = 200
Sample size = 20
n = 20
PROBABILITY IS: 3.32169e-07

This works out to about 1 chance in 3 million- an extremely unlikely outcome. Even though we look at a small number of marbles- only 20 - the improbability becomes enormous. This shows why these types of problems are sometimes not completely obvious, at first glance.

This is not to say that the outcome cannot *ever* happen. It does indeed happen. This would be almost exactly the rate at which you would expect to find it happening, if you repeated the experiment an infinite number of times- roughly one time out every 3 million attempts.

What if you are trying to analyze a situation where there are multiple possible explanations for getting the result that you did, besides simply one of random chance? Suppose, for example, that the experimenter forgot the importance of evenly mixing the marbles in the jar. Instead, 100 black marbles were poured on the bottom, and 100 white marbles poured on top. This might guarantee a near 100% chance that you could pull a handful of all-white marbles, instead of the rightful probability of near-zero.

Most of the preceding problems have involved an equal number of marbles of one color an another. You can also have imbalances in the numbers of marbles of each type in the jar. If you have 9999 white marbles, and only 1 black, then draw one marble, the chance of getting the black marble is very small (1 in 10000).

Our program is geared for the general case: X objects of one type, Y objects of another type, drawing a sample of Z objects, then computing the probability of getting "n" or more of the "X" type in the sample of Z. You fill in the values of X, Y, Z, and n, for any problem of this sort.

Suppose that this situation were one where you were trying to evaluate the safety of a food additive, to make sure that it was not a carcinogen (cancer-causing). Say, that you have studied a large population of X animals, and that in a given year, you normally see Y animals that get cancer (maybe this would be a flat percentage, such as 1% of X). Now, you study a smaller group of Z animals, using the food additive. You find that out of these, you have "n" that get cancer in a year. You would expect for "n" to be roughly equal to Z * (Y/X). You expect it to be ***slightly*** different than this, because of random chance variation, but you do not expect it to be ***greatly*** different. You can compute the probability for the difference that you see, just as you can for a problem in drawing marbles from a jar.

Say that the odds against a higher cancer rate being attributable to random chance is only 1 in a million. Do you approve the food additive (or in our case, the vaccine)? Of course not. There is indeed still a chance that the higher rate is only random, but you would not take that kind of chance.

What constitutes "statistical significance" is not something that has a completely hard-and-fast rule. A common convention used in many evaluations is a level of "1 in 1000". Other, less demanding problems might test for significance at, say, the 1% level (1 in 100), or the 5% level (1 in 20).

The desired level of significance is often chosen based on the needs of the problem under consideration. Since our problem is a critical one of vaccine safety, we are more than meeting a reasonable definition of "statistical significance".

If you draw a graph of probabilities for different outcomes, for these types of problems, you tend to get a "bell-shaped" curve. The center of the curve is the most expected result (such as 50% white marbles drawn in a sample from a jar with a 50-50 white and black mix). The edges of the curve as the least likely results (such as all-white marbles).

As your population size and sample size grow (i.e., more marbles in the jar and bigger handfuls taken out), the odds against getting any one ***particular*** outcome (such as getting exactly 1,203 white marbles and 2,459 black marbles in a giant handful) becomes extremely small. This type of "improbable" result is nothing unusual or suspicious, because there are many, many such uninteresting, particular mixes that are equally possible.

Our program is not computing odds for single outcomes of this sort. It is computing a whole range of combined possible outcomes ("n or more"). This is in effect an "area under the curve" for the bell-shaped probability curve. That is why the results are meaningful, regardless of the population and sample sizes, or the shape of the curve.

As mentioned earlier, the extremely small probability values computed in our analysis point to the fact that our sample sizes are large enough to be significant. Smaller populations and samples tend in general to yield higher probability values.

For example, the odds of getting 5 heads out of 6 coin tosses is about 10%, which is perfectly feasible. The odds of getting 500 heads out of 600 coins tosses is about 3 times (10 to the -65th power). A trillionth would be 10 to the -12th power, so this is an unimaginably small possibility.

The ***ratio*** of "head" outcomes to the total coin-toss trials is the same in both cases, at 5:6.

Yet the probabilities are vastly different. Six coins tosses are clearly not enough to guarantee that the outcome of 5 "heads" was not by random chance. Six hundred coins tosses are far more than enough.

The same phenomena apply to the 799 men in the SFMS and the 359 in the SFHBVCS.

The computed probability is the real probability, and is fully meaningful, so long as the data used to make the calculation are reliable.

If the formula or program seem complicated, that is the only real illusion. It is enough to realize that this is a common type of problem, found in nearly any statistics text. It would be pointless to attempt to create a mirage, in a document that will subject to the criticism of experts.