

DISCUSSION

Stephen E. Fienberg
Carnegie Mellon University
Pittsburgh, PA 15213

1. Prologue.

What do we know about the prevalence of HIV infection in the United State? What do we need to know in order to make wise policy decisions and effective planning? These are the key issues that serve as a backdrop to the two papers presented in this session.

Our current knowledge of the extent and consequences of the AIDS epidemic has come chiefly from the Centers for Disease Control (CDC), and their national disease reporting apparatus as well as the so-called CDC family of surveys. The former produces a tally of known and reported cases of the AIDS complex of diseases while the latter is an ad hoc collection of ongoing observational studies. Current national estimates of HIV prevalence are then derived from these data using methods such as back-calculation or by gluing together disparate sources and extrapolating. In other words, today we do not know to any reasonable degree of accuracy what the prevalence of HIV is, nor how infection by the virus is distributed over the population. At best, we simply have some strong clues. At worst, we have biased and misleading information.

Let me now turn to the two papers and begin with a summary of my impressions of the them and of the extensive statistical effort that they describe. I have followed the work on the pretest in various ways over the past two years: first as a member of the Allegheny County Community Advisory Committee for the Pilot survey and then as an advisor to CDC and NCHS on the Dallas Pretest.

The results of the pilot and pretest programs are very impressive and I am pleased to be able to publicly applaud the

thorough planning effort by the staff at NCHS and at RTI. I think they have done an excellent job and they have convinced me of the wisdom in proceeding with plans for a full scale national survey.

Having said this I should note that there are some residual problems that RTI and NCHS will need to continue to worry about as we look towards a full-scale survey, especially those problems that relate to the adjustments for nonresponders. Moreover, I would do certain analyses and the reporting of certain results in a somewhat different fashion. These comments should not be misconstrued, however, as a criticism of the RTI effort but rather should be interpreted as part of the orderly development of scientific studies. After all a discussant of papers at an ASA session is supposed to make some critical comments!

The remainder of my discussion will take the form of answers to a subset questions raised earlier this summer in a review of the outcomes of the NHSS pretest project. In this discussion I will not actually distinguish between the two papers since I view them as parts of a whole.

2. Questions and Answers.

1. *Is the response achieved in the feasibility study sufficient to support using a household survey approach?* In my view, yes. To date we have had only observational data and statistical manipulations like those associated with the method of back-calculation to rely upon for getting a handle on the prevalence of HIV infection. The only other surveys available on a national basis are those in the CDC Family of Surveys which consist of convenience samples of special populations that are clearly not representative of the major segments of the

US population. I have asked myself several times whether there were any viable alternatives to a national household sample for baseline data at the moment, and the answer is always "no". On the other hand I do think that expanding the frame for the survey in one or more ways would be of great value and thus I would urge that this matter be explored further.

2. *Did the feasibility study in dictate that non-response and other biases . . . could be accurately evaluated and controlled for? . . .* Again, my answer is yes. The NCHS/RTI group did a first-rate job of analyzing the Dallas data and in developing preliminary methods for controlling and adjusting for nonresponse bias. Their report to NCHS and the present pair of papers outline how this work should be extended as we move to a national survey. I would add that more work could be done on this problem but I am confident that the kinds of methods used to date can ultimately be adapted to handle adjustments in the NHSS. For example the notion of fitting a linear logistic component to the highly nonlinear frequency categories seems strange at best, even though it has moderately good predictive power.

3. *Did the feasibility study demonstrate that a national survey can collect accurate information on risk behavior in the US?* This is a tricky question to answer. We have no benchmark by which we can judge the results of the pilot and pretest surveys. What we do know is that a substantial number of people are willing to report high risk behavior. What we do not know is whether these reports of such behavior are accurate nor whether those not reporting such behaviors are telling the truth. My somewhat informed guess is that reported high risk behaviors are relatively accurate except for frequencies of occurrence, and that the number of people reporting such behaviors is a modest underestimate of actual behavior. The results for those who are HIV positive give modest support for such a guess.

4. *Did the feasibility study demonstrate that similar data collected nationally would be scientifically useful for:*

- *estimating the prevalence of HIV infection in the US?* Yes.

- *estimating the level of HIV-associated risk behaviors?* Yes, subject to some underreporting as noted above.

- *estimating the prevalence of HIV infection by reported risk behaviors?* Yes, although if NCHS is interested in such prevalence rates for demographic subpopulations some methods of "smoothing" the data will need to be employed.

- *validating results from back-calculation and other statistical and epidemiological models for HIV prevalence?* Based on my reading of the statistical literature, I believe that the methods employed in back-calculation and other forms of projection from observational studies and reported AIDS cases are fraught with problems and problematic assumptions. This is not a criticism of that literature, which I consider to be an excellent example of ingenious statistical methods developed to deal with a very tricky statistical problem. Rather this is an observation on the difficulty of the task. As a consequence, I was impressed that we could actually begin to reconcile the prevalence rates in Dallas with those produce by the cruder methodology and data available for back-calculation. I would hope to see a closing of the gap for national estimates as both the NHSS and the back-calculation methods get refined. In fact, one shouldn't think in terms of validating back-calculation results which surly are incorrect. Rather, one should ask how the survey results will help researchers refine back-calculation and other methods to make them more accurate than they currently are.

5. *Would the indirect adjustment of HIV prevalence demonstrated in the pretest be a valid and useful technique in a national survey?* Yes, subject of course to further

refinements and embellishment as suggested in the second of the two papers and as I noted previously.

6. Is the direct assessment study or some other statistical verification study necessary for conducting a national survey? Faced with a choice between a national HIV survey without such a study and no national survey I would opt for the former because I think that the current methodology is relatively good and that we as a nation are desperate for some sort of population-based prevalence estimates. On the other hand doing some type of direct or even indirect assessment is virtually always advisable in national surveys of this sort and that it would be wise to incorporate them into a NHSS.

7. Did the pretest evince sufficient evidence that an anonymous survey can be conducted in a manner to effectively convince respondents that their privacy is being protected? As a close observer of the process in Allegheny County I remain convinced that the pilot study could have been carried out without the switch from strict confidentiality to anonymity. As the only technical-trained individual on the Allegheny County Community Advisory Committee, I know that virtually none of the others understood the distinction between

confidentiality and anonymity. Thus I believe we have not yet had any direct test that would support the statement in one of the papers to the effect that "It is not feasible to conduct a NHSS without . . . assur[ing] that the data collected are anonymous." This statement is speculation; it is more of a political and social observation than a statistical one since it is based on a single piece of data, the fiasco that led to the cancellation of the Washington, D.C. pilot survey. There is no question in my mind that NCHS and RTI went overboard to ensure the anonymity as well as the perception of anonymity, and for this they are to be congratulated.

3. Epilogue.

AIDS remains a major public policy and health problem in the United States. The annual expenditure of billions of dollars a year on research supported by NIH and the many millions of dollars on data collection by CDC have produced neither a cure nor an accurate estimate of the dimension of the problems we face as a nation in the future.

The design of the NHSS is a good one and it can go a long way towards providing needed data for policy purposes. I think we need to get on with the job. Let's get the survey funded and into the field.