THE UNITED STATES IMMUNIZATION SURVEY: A REVIEW OF ITS IMPORTANCE AND SUGGESTIONS FOR IMPROVEMENT

Judy A. Bean, Leon F. Burmeister, and Charles L. Paule, University of Iowa

1. Background

The Current Population Survey (CPS) is conducted monthly by the United States Bureau of the Census with the primary objective of describing the current economic status of the country. The CPS has provided unemployment statistics from the late 1930's to the present time.

In recent years supplemental information has been collected in addition to the basic employment statistics. These surveys include immunization against selected diseases, school enrollment, voting in national elections, marital history, fertility, income and other areas of interest (1).

The United States Immunization Survey (USIS) is conducted every September in conjunction with the CPS. This is accomplished by attaching to the CPS questionnaire a supplemental instrument, which primarily concerns the immunization status of children less than 14 years old for the diseases influenza; polio; diphtheria, tetanus and whooping cough; measles; rubella and mumps, as well as the history of measles, rubella and chicken pox infections.

The results of recent United States Immunization Surveys have caused concern. The estimated immunization rates for several important diseases have not increased in recent years. In fact, the immunization rate for polio has apparently decreased, resulting in several warnings of possible epidemics (2,3). Since the impact of such outbreaks could be extreme, concern has been expressed not only about the possibility of epidemics, but about the accuracy of the USIS. That is, is the alarm that has been sounded possibly false, or do such statistics describe a real problem?

One cause of suspicion that the USIS rates are too low is the fact that several other sources of immunization data indicate higher immunization rates. However, such data are descriptions of local or state immunization status, as opposed to regional and national immunization rates.

An additional source that could be interpreted as implying the rates indicated by the USIS are too low is the amount of vaccine distributed by drug companies. For example, 3.5 million children enter the age for measles vaccination each year, yet 8 million measles vaccine doses are apparently distributed. Possible reasons for the discrepancy include: some previously immunized children are re-vaccinated, many distributed doses are never utilized, and the immunization rate for measles is higher than indicated by the USIS.

Obviously, it is very important to carefully assess the validity of the immunization rates estimated by the USIS. If the potential for epidemics in some of the common childhood diseases exists, health professionals must have reliable indicators of such potential. However, the USIS is a national study, and its smallest areas of estimation are multi-state regions. Rates estimated for large areas cannot be expected to predict the occurrence of outbreaks in particular cities or counties. Not only is a thorough investigation of the accuracy of the USIS necessary, it is also important to carefully examine the purposes of this national sample of immunization status of children.

2. Methodology

The methodology utilized to estimate the validity of the USIS can be classified into two areas: an evaluation of the sample survey design selected to estimate employment rather than immunization levels, supplemented by an evaluation of problems such as vaccine take rates that are peculiar to immunization studies. The statistical rather than the biological aspects of the USIS are the primary concern of this paper. Such concerns include the survey design, the estimation procedures employed and interviewer methods and training.

The results from the above approaches are greatly enhanced by comparisons of the USIS estimates to independent studies done at state levels. In addition the changes in immunization levels estimated by the USIS from one year to another when supplemented by historical information relative to vaccine emphasis and known outbreaks are extremely interesting.

3. Sample Design of the Current Population Survey

The CPS is a complex multistage probability sample design employing the concept of rotation panels. Approximately 47,000 households are selected for interview each month. From the data collected, estimates not only of unemployment for the nation, but for geographical regions, specified residence categories and certain age-sex-race classes are also produced.

The basic sampling units are the nation's 3,146 counties which are grouped into 1,931 primary units (PSU's) with each of the 237 Standard Metropolitan Statistical Areas constituting a single PSU. The PSU's are stratified using the following variables: geographical region, population growth, population density, proportion nonwhite, principal industry, number of farms, retail sales per capita and relative number of hotels and motels. To form strata approximately equal in population size, the one hundred fifty-six largest PSU's each form a separate stratum with the remaining 1,775 PSU's classified into 220 strata.

Within each PSU, the Bureau of the Census has created enumeration districts (ED's) which contain, roughly, 300 households. Within ED's the households are further subdivided into clusters of four called segments. The households within a segment are geographically contiguous.

After a sample is chosen, it is divided into eight groups in such a way that each is a miniature probability sample of the United States. The segments in each resulting rotation group are interviewed for four months, not interviewed for the next eight months and then interviewed again for four months. After the last four months, the segments are eliminated from the sample. Thus, 75% of the sample is the same from

month to month and 50% from year to year.

Each month 47,000 households are designated to be interviewed in the 461 PSU's. Because of nonresponse, approximately 1,700 of the households are noninterviewed; in addition, 7,500 households are found to be vacant or not enumerated for other reasons.

The final published estimate of a characteristic is a composite estimate based on a weighted average of two estimates for the current month. The first estimate results from the sample data collected in the current month. The second estimate is derived by adding to the preceding month's composite estimate an estimate of change. This estimate of change is based on the sample data which is common to the two months (75% overlap). Each estimate receives a weight of 0.5.

The details of the Current Population Survey such as the steps taken in drawing the sample and the weights used at various stages of estimation are described by Thompson and Shapiro(1).

4. United States Immunization Survey

Questions concerning immunization and infection status for certain diseases added to the September CPS questionnaire are asked principally for children less than 14 years old. These data are used to provide estimates of immunization levels for the nation, geographical divisions, certain residence classes and specified age-race groups.

Because of cost factors, the set of immunization questions is only asked in six of the eight rotation groups. The two omitted are the panel entering for the first time and the one reentering after being out for eight months. In the six groups, approximately 35,500 households are designated to be interviewed; data for only 28,000 are collected since 1,500 households are nonrespondents and 6,000 others are found to be vacant.

A written description of exactly how the final estimates are produced apparently does not exist; however, the Bureau of the Census has indicated that a composite estimate incorporating an estimate of change based on the common sample data from year to year is not used.

5. Results of the Review of the USIS

5.1 Use of Rotation Panels:

The use of panel studies can improve estimates depending upon the nature of the desired estimates in addition to characteristics of the population being studied. To evaluate the degree of matching utilized in the USIS, it is necessary to know the correlation of responses on successive occasions. There are no data available to answer this but a reasonable assumption is that there is a strong positive correlation between successive responses on immunization questions. Therefore, the sample design is not optimal, Cochran (4), if the primary interest is in current estimates because the USIS utilizes an overlap of year-to-year.

However, users of the USIS are concerned with the estimation of trends as well as current estimates. When considered separately, the trend estimation would necessitate complete overlap, opposed to the current estimation, which would indicate little overlap, assuming high correlation between successive responses. Thus the 50% overlap appears to be a reasonable compromise if both estimates are considered of equal importance.

5.2 Design Difficulties of Panels:

One of the limitations of repeated sampling is the "conditioning" of the respondents. Waksberg and Pearl (5) review the effect of conditioning on the CPS. Although the paper is primarily concerned about the effects of conditioning on the labor force statistics, it is helpful to review the conclusions and speculate about possible conditioning effects on the USIS. For example, the conditioning of respondents could cause an abnormal number of parents to have children vaccinated during the year between interviews.

Waksberg and Pearl note that interviewers are more likely to classify a housing unit as vacant as the number of interviews increases, and that such an increase in vacancy occurs primarily for rented units. This increase in the declaration of vacancy of rented units possibly has an effect on the USIS. Assume that renters are less affluent than home owners and consequently are less likely to have their children immunized. Then if a disproportionate number of rented households are incorrectly declared vacant and are incorrectly omitted from the immunization survey, the estimated percentage of immunized children would be too high.

Bailar (6) shows that not only is the panel study subject to bias which results from conditioning, but the estimate of the variance is likely to be biased upward. Both the ratio estimator and the composite estimator of the monthly level are biased, although the estimators of month-to-month change are essentially unbiased. The estimators of the variances of both the estimates of monthy level and month-to-month change are also biased.

5.3. Nonresponse:

Thompson and Shapiro (1) describe three types of nonresponse in the CPS. These are:

(1) Undercoverage

The CPS coverage is measured by the ratio of the sample population 14 years and over to the independent estimated totals. The ratio is almost always below one meaning that some subgroups of the population are being missed every month.

(2) Noninterview

Each month a portion of the sample households that are occupied are not interviewed for differing reasons.

(3) Nonresponse

This refers to items that are left blank on the completed questionnaires.

Technical Paper No 19 (7) describes an experiment conducted in the October, 1966 and June, 1967 reinterview programs. The reinterview procedure should have indicated the causes of undercoverage but in practice this was not the case. This intensive coverage check was carried out to determine the reasons for the undercoverage which would provide a basis for improving the training of the interviewers. One important finding was that a significant source of coverage loss was due to misclassifying noninterviews.

To know the effect on the USIS, the composition of these missed households has to be determined. Palmer (8) describes an intensive follow-up of

nonresponse households but even this effort resulted in only about half of the households being interviewed. The most significant finding was that approximately 60% of the nonresponse households were either one person or two person households. This would seem to indicate that the households being missed would not affect USIS estimates because there would be few children under 14 years old in these houses.

5.4 Instrument:

The questionnaire employed is another potential source of bias. Apparently no pretesting of the questions has been done since the middle 1960's. The only possible responses for the majority of the questions are "yes" or "no" which allows the possibility of bias due to omission of other alternative responses. The interviewer is instructed to probe if the respondent's initial response is "I don't know". There is no way to tabulate how many people are forced to respond when the true response is "don't know", given the fact that only blanks have been tabulated as "unknown".

5.5 Telephone Interviews:

A personal visit is always required for the first and fifth months of the CPS. During the second month the first attempt to interview a household should also be a personal visit. However, if no one is at home, a telephone call-back interview may be used. The telephone is used by the interviewer during the third, fourth, sixth, seventh and eighth months if an agreement is made in a personal interview to subsequently use the telephone. Call-back telephone interviews are used when a household which is normally interviewed personally is found to have no one home. If contact is made over the telephone, the interview may be conducted. When making regular telephone interviews, if no one answers despite repeated calls, the interview is designated as a "noninterview."

Since the USIS is conducted in the second, third, fourth, sixth, seventh, and eighth rotation groups, one could conclude that a large proportion of the interviews are conducted over the telephone. Detailed study of the effect of telephone interviews should be completed because telephone interviews form a potentially large proportion of the interviews conducted for the USIS.

6. Findings in Relation to Certain Uses of the USIS Data

In evaluating the accuracy of the USIS it is necessary to consider the current uses of the data since it is possible that results are sufficiently accurate for some purposes but not for others.

6.1 Estimation of Change in Immunization Rates Over Time:

Estimation of change in the rates over time is the most common context in which the USIS data are quoted and used. An example is the expressed fear that epidemics of poliomyelitis may reappear since rates of immunization have declined over recent years.

Since estimation of time trends, rather than of rates at one point in time, is considered more important, a number of possible methodologic errors would not seriously affect these estimates provided the magnitude and direction of the error remain constant over time. Problems such as the lack of "don't know" responses may introduce a bias, but this would not seriously affect trends under conditions in which the survey instrument and type and training of interviewers remain constant.

A more pertinent criticism is the failure of USIS to make use of a composite estimator. The estimator currently being employed fails to take advantage of the data collected the previous year on the same respondents. The utilization of existing overlapping panels would improve the sensitivity of changes from year to year.

The increasing use of measles, mumps and rubella vaccines in combination rather than as individual vaccines may affect the estimates. Since no data are available to indicate the direction of the error, both under and over reporting of individual vaccines could be logically suggested. No evidence could be found that changes were made in either the survey instrument or interviewer training to take this modification into account. A similar situation has occurred with poliomyelitis vaccine, with a shift having been made from separate administration of types I, II and III to a combined form. Use of the combination has been virtually routine for several years; therefore, this change in form of vaccine should not affect time trends over recent years.

6.2 Estimation of Current Immunization Levels:

Current estimates without relation to change over time may have value in locating needs for additional immunization effort according to vaccine type or by broad geographic and socioeconomic groupings. Because of the limited sample size, estimates are only produced for regions of the United States, by age groups, for Standard Metropolitan Statistical Area (SMSA) components by age groups, for SMSA components by poverty status by age groups and for single year of age for the entire United States. Estimates for many other important demographic subgroups such as states or counties are not included in the design. The cells presented are useful only at the national and regional level.

6.3 Aid in Planning Local Immunization Campaigns:

As stated in the preceding section, the USIS sample sizes have been designed primarily to provide data significant only at a gross regional level. While representative communities throughout the country have been included in the sample, there are insufficient numbers from any single metropolitan area or from any single state to allow accurate estimates of immunization rates from defined local areas. Consequently the USIS data are not useful for local planning.

6.4 Determination of Levels of Immunity:

Immunity may be induced by either natural infection or immunization. Since the disease history data obtained by the USIS are limited, the data are representative only of vaccine-induced immunity. As reservoirs of natural infection decrease, immunization history will become more indicative of immune levels. At the present time a significant amount of natural infection with the diseases included in the survey still occurs, so rates of true immunity as assessed only by artificial immunization are low.

History of immunization is itself subject to

biological limitations. Significant failures to respond to immunization have been noted under conditions of normal use for most of the vaccines under review and must be taken into account. In addition the duration of protection afforded by the vaccines has not been completely resolved and may become a particularly significant problem in the future if the boostering effect of subclinical infection declines.

6.5 Prediction of Disease Outbreaks:

The ability of the USIS to predict disease outbreaks is limited by: a) as outlined above, true immunity is not directly measured; b) limited sample sizes preclude statistical significance at local community levels, and c) even if true immune levels were precisely known, the exact levels of susceptibility at which large scale epidemics may propagate have not been determined. Anticipation of epidemics may only be crudely inferred from USIS data.

On the other hand, if additional outbreaks of disease occur, the distribution of cases can be roughly predicted by USIS data. This has been most evident for measles, which has shown a socioeconomic distribution of cases consistent with changing patterns of vaccine utilization as measured by USIS. To a certain extent, the USIS has been successful in identifying large groups at particular risk of disease.

7. Comparison of USIS Measles Data to Other Sources of Immunization Estimates

Figure 1 describes a family of birth cohort curves of rates of history of measles infection. If the estimates were absolutely accurate, the cohort curves would either remain flat if no new cases occur or would rise if new cases occur, but under no conditions would they show a decline. The declines, therefore, indicate error in the estimation procedure and/or the survey. The error could result from sampling variation since only 50 per cent of children are true cohorts in the sense that they are the same persons followed from the previous year. Simple chance variation in sampling of new populations could thus lead to discrepancies. The other type of error could be the variations of response to the question asked in the survey. If this were the case, it would be necessary to postulate that such response errors are not constant, but vary in magnitude from one year to the next. Similarly rises could be due to sampling or response errors, but may also represent true increases. The groups born between 1960 and 1973 generally show a gradually increasing curve which is expected. However, the cohorts from 1956 to 1959 show a decline in the percentage of children with measles infection. This may be explained by the following theory. As each year passes in a cohort group, a certain percentage of children is likely to contract measles. This percentage would get progressively smaller over time. Simultaneously, there is a lack of recall in a certain percentage of those who have had measles due to the passage of time between the date of the infection and the time of the interview. This percentage will increase with time. At some point the percentage of recall mistakes overtakes the percentage of children contracting new infections which will cause a downward trend in the curve.

Figure 2 compares USIS estimates of the yearly number of measles cases with reported cases from the Morbidity and Mortality Weekly Reports (MMWR). The MMWR figure is based on an accumulation of local case reports. In a 1973 Measles Surveillance Report the reporting of cases was estimated to be less than 10%; therefore, in a crude attempt to make the MMWR more accurate, the MMWR figures are presented ten times their actual size(MMWR*).

Both sets of data show the same general time trends, a downward trend from 1966, an increase from 1968-71, and a subsequent decline. There are two major differences. The USIS information would cause one to believe that 1970 was a peak year for measles, while the MMWR shows a peak in 1971. The second difference is the marked discrepancy between the number of cases, even when a correction factor was applied to the MMWR data. A slight increase during 1973 shown by the USIS data was not reflected by MMWR reports.

A comparison of data obtained from the Center for Disease Control's Biologic Surveillance Reports with USIS data on estimated yearly vaccination totals is given in Figure 3. The magnitude of the differences between the two curves may be due to the fact that the Biologics Surveillance is a summary of vaccines manufactured and not doses administered. An approximate similarity between the two curves can be observed. When biologics distribution was at a low in 1970, there was a corresponding low in the number of children receiving immunizations that year. This was followed by an increase in 1971 of both distribution and vaccination. When distribution dropped in 1973 there was a corresponding drop in reported immunizations.

8. Conclusions and Recommendations

The United States Immunization Survey (USIS) is conducted every September in conjunction with the Current Population Survey (CPS). The importance of the USIS cannot be over-emphasized since the possibilities of outbreaks of diseases such as polio must be identified and then minimized to maintain the present level of our country's health. However, the USIS is a nationwide study and is not of sufficient size to accurately estimate immunization levels for small areas.

The USIS includes additional potential sources of error such as nonresponse bias, apparently forced "yes" or "no" answers, failure to validate immunization questions, lack of data on effect of conditioning, diagnostic difficulty and both educational and cultural biases in recall.

However, the importance of the USIS should not be judged in view of these difficulties, but with respect to its realistic uses. It is not reasonable to expect precise estimates of immunization at a local level; however, the USIS is extremely helpful in determining changes in immunization rates from year to year in addition to identifying those types of people most susceptible to potential epidemics. Although the USIS has aspects in need of improvement, these two uses of its results justify its current form of existence.

References

1. Thompson, Marvin M. and Shapiro, Gary, "The

Current Population Survey: An Overview," Annals of Economic and Social Measurement, 2 (1973), 105-129.

2. ———, "Immunization vs. Complacency: Are We Ready for the Challenge?," Journal of the American Medical Association, 229 (September 1975), 1557-1558.

3. ———, "Diseases That Should Not Happen," Medical World News, 15(September 1974), 33-42.

4. Cochran, William G., Sampling Techniques, John Wiley and Sons, 1963.

5. Waksberg, Joseph and Pearl, Robert B., "The Effects of Repeated Household Interviews in the Current Population Survey," 47th National Conference of the American Marketing Association, Dallas, Texas, June 17, 1964.

6. Bailar, Barbara, A., "The Effects of Rotation Group Bias on Estimates from Panel Surveys," Journal of the American Statistical Association, 70(March 1975), 23-30.

7. U.S. Department of Commerce, Bureau of the Census, The Current Population Survey Reinterview Program, January, 1961, through December, 1966, Technical Paper 19, U.S. Government Printing Office, Washington, D.C., 1968.

8. Palmer, Susan, "On the Character and Influence of Nonresponse in the Current Population Survey," Proceedings of the Social Statistics Section, American Statistical Association, (1967), 73-80.

ACKNOWLEDGMENT

This work was partially supported by the Center for Disease Control, Contract No.: CDC 21-74-545.

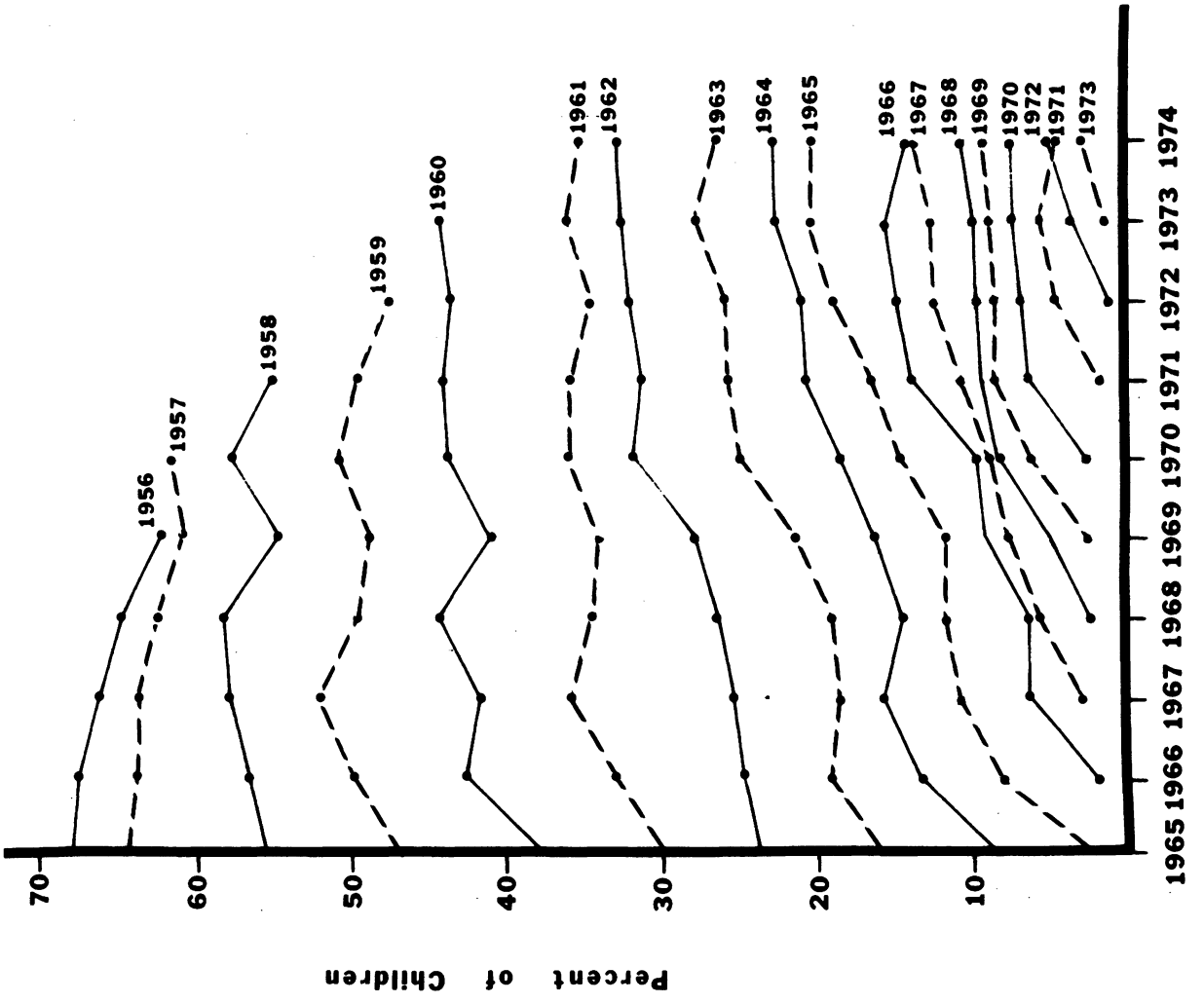

Figure 1. The estimated percent of children with history of measles infection by birth cohort, USIS, 1965-1974. Each line is a birth cohort group followed by the year of birth.

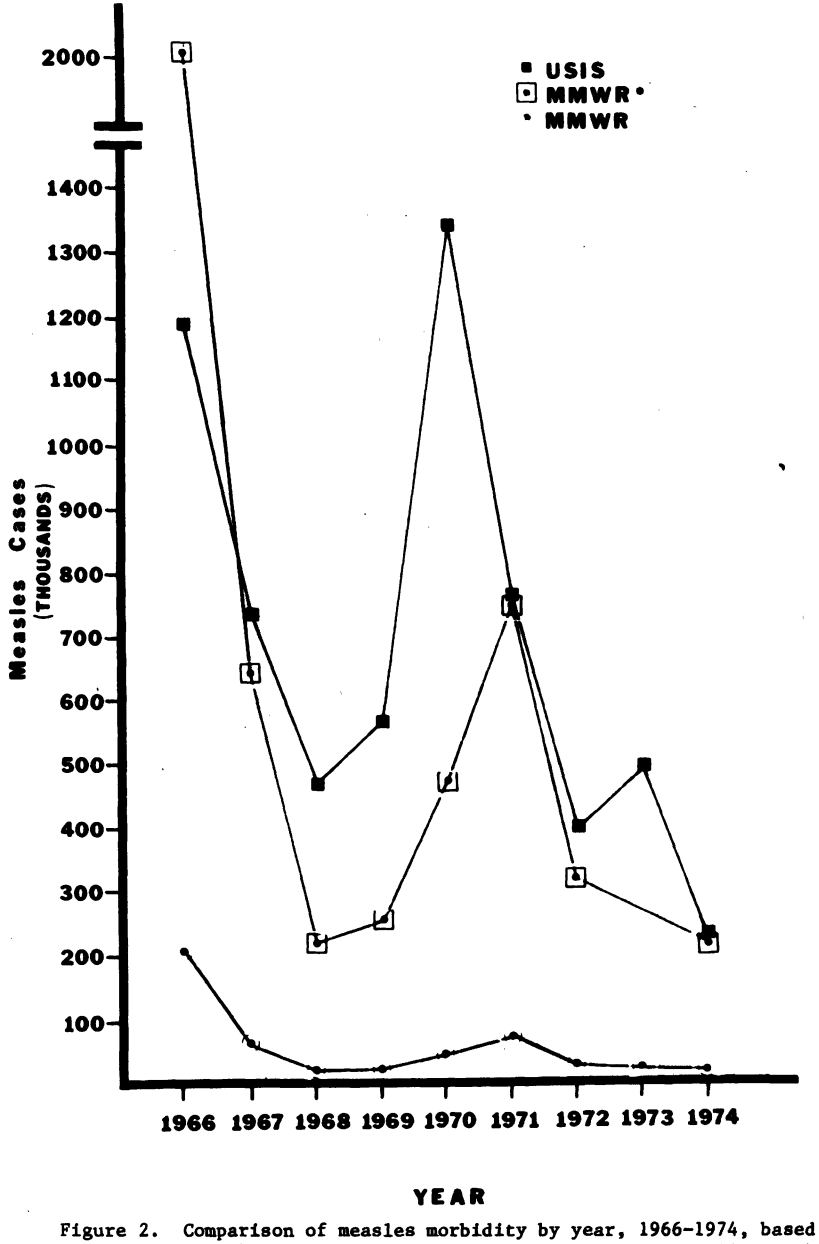

Figure 2. Comparison of measles morbidity by year, 1966-1974, based on 1) reported cases by the Morbidity and Mortality Weekly Reports (MMWR), 2) adjusted reported cases (Morbidity and Mortality Reports multiplied by ten) (MMWR*), 3) estimates derived from the United States Immunization Survey (USIS).

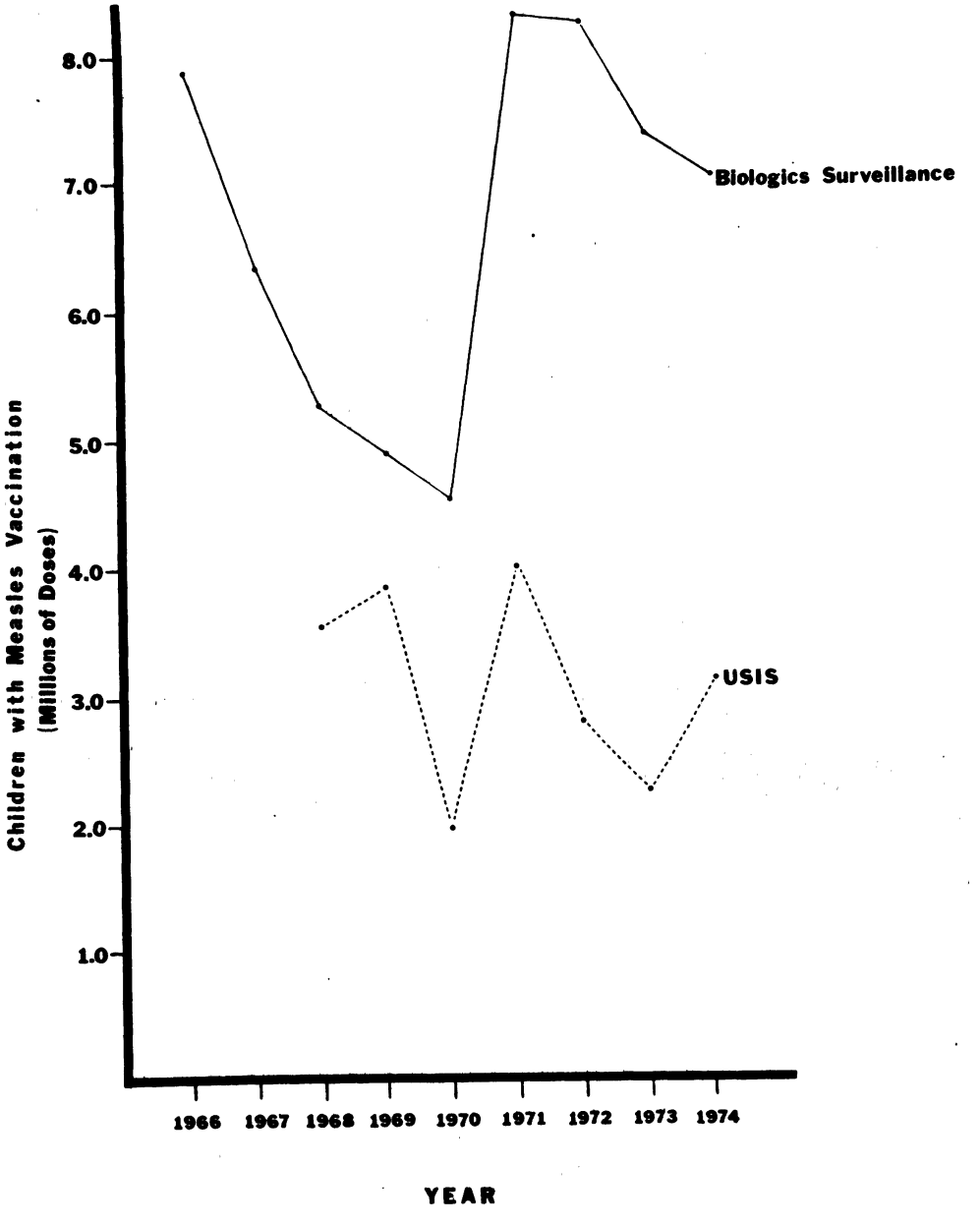

Figure 3. The comparison by year of doses of measles vaccine distributed to the derived estimate of the number of children with measles vaccination in the past year, 1966-1974.