Archive Copy

UM-HSRI-SA-75-18

The National Electronic Injury Surveillance System and Bicycle-Associated Accidents

Jairus D. Flora, Principal Investigator
Richard J. Kaplan
Elizabeth Margoshes
Peter D. Ward

October 1975

Highway Safety Research Institute/University of Michigan

Final Report to
Bicycle Manufacturers Association

THE NATIONAL ELECTRONIC INJURY SURVEILLANCE SYSTEM AND BICYCLE-ASSOCIATED ACCIDENTS

FINAL REPORT

to

Bicycle Manufacturers Association

31 October 1975

Jairus D. Flora, Principal Investigator
Richard J. Kaplan
Elizabeth Margoshes
Peter D. Ward

Prepared By

Highway Safety Research Institute
The University of Michigan
Huron Parkway & Baxter Rd.
Ann Arbor, Michigan 48105

Technical Report Documentation Page

1. Report No.	2. Government Accession No.	3. Recipient's Catalog No.
HSRI-SA-75-18		

4. Title and Subtitle	5. Report Date
The National Electronic Injury Surveillance System and Bicycle-Associated Accidents	31 October 1975
	6. Performing Organization Code
7. Author(s) J.D. Flora, R.J. Kaplan, E. Margoshes, and P.D. Ward	8. Performing Organization Report No. HSRI-SA-75-18
9. Performing Organization Name and Address Highway Safety Research Institute The University of Michigan Ann Arbor, Michigan	10. Work Unit No. (TRAIS)
	11. Contract or Grant No.
	13. Type of Report and Period Covered Final Report
12. Sponsoring Agency Name and Address Bicycle Manufacturers Association	14. Sponsoring Agency Code

15. Supplementary Notes

16. Abstract

The National Electronic Injury Surveillance System used by the Consumer Product Safety Commission has produced data which lead to national estimates of product-associated injuries. This report presents the results of a study of the methods of data collection used by the NEISS and the uses made of such data. Two major areas for improvement of the NEISS are suggested: changing the design of the hospital sample and instituting primary data collection rather than sampling existing hospital records. Product comparisons implied by the "age adjusted frequency severity index" calculated by the CPSC are found to be quite sensitive to the particular injury severity scale used and to the particular age weighting factor used. Alternatives to the numerical scales currently used are suggested. An analysis of bicycle-associated injuries based on data from the NEISS is presented which shows a nearly stable total after adjustment for seasonal effects. Misinterpretations of the data which have appeared in various publications are noted and cautions against the misuse of the NEISS data are presented.

17. Key Words	18. Distribution Statement

19. Security Classif. (of this report)	20. Security Classif. (of this page)	21. No. of Pages	22. Price

Form DOT F 1700.7 (8-72) Reproduction of completed page authorized

CONTENTS

ABSTRACT

The National Electronic Injury Surveillance System used by the Consumer Product Safety Commission has produced data which lead to national estimates of product-associated injuries. This report presents the results of a study of the methods of data collection used by the NEISS and the uses made of such data. Two major areas for improvement of the NEISS are suggested: changing the design of the hospital sample and instituting primary data collection rather than sampling existing hospital records. Product comparisons implied by the "age adjusted frequency severity index" calculated by the CPSC are found to be quite sensitive to the particular injury severity scale used and to the particular age weighting factor used. Alternatives to the numerical scales currently used are suggested. An analysis of bicycle-associated injuries based on data from the NEISS is presented which shows a nearly stable total after adjustment for seasonal effects. Misinterpretations of the data which have appeared in various publications are noted and cautions against the misuse of the NEISS data are presented.

I. INTRODUCTION

The Consumer Product Safety Commission was activated on May 14, 1973 as a result of implementing the Consumer Product Safety Act (PL 92.573) signed into law on October 27, 1972. Upon its formation responsibility to administer several protection acts (Flammable Fabrics Act, Federal Hazardous Substance Act, Poison Prevention Packages Act, and Refrigerate Safety Act) was transferred to the Commission. The commision sought a single priority system to evaluate the risks of the more than 10,000 consumer products.

The Bureau of Epidemiology has been given the responsibility by the Commission of determining the essential facts about product-related accidents throughout the nation. This task includes making a determination of how the Commission should allocate its financial resources in pursuing products which present substantial or unusual risk to the consumer. To aid them in this task, the Commission staff has at its disposal a data acquisition system known as the National Electronic Injury Surveillance System (NEISS). The NEISS had been developed by the Bureau of Product Safety of the Food and Drug Administration and has been in continuous operation since July, 1972.

Data come to the NEISS from a statistical sample of hospital emergency rooms selected from a population of over 5,000 such facilities in the United States. Data from these

119 emergency rooms are coded and tabulated by computer to show a daily sample of product associated injuries. From tabulation of the NEISS data, cases are selected for in-depth investigation to determine details of how a product was involved and the course of events which resulted in the injury.

The NEISS data can also be used to estimate national totals of various types of product-associated injuries. For example, national totals of injuries by product, age, sex, injury, etc., may be estimated. Combinations of the frequency and severity of injury and the age of injured individuals were used by the Commission in arriving at the Consumer Product Hazard Index, which the Commission first published in 1973.

Bicycling has experienced a marked rise in popularity in the last few years, both as a form of recreation or sport and as a means of personal transportation. The latter use received additional impetus from the recent energy crisis and the resulting gasoline shortage. The growth in popularity for bicycling has resulted in an increased use of bicycles by adults and a spurt in bicycles sales.

Nearly coincident with the upsurge in bicycle popularity has been the development of the National Electronic Injury Surveillance System (NEISS) by the Consumer Product Safety Commission. Preliminary CPSC analyses of the data produced by the NEISS were directed to determining a rational order in which to investigate various consumer products to determine if they posed an unreasonable risk. The resultant ordering of products has been called--rather inaccurately--The Consumer Product Hazard Index and has been widely publicized by the media. Bicycles turned out to be at the top of the ordering of products. This has led to some concern both by the consumers--persons returning to bicycling or considering bicycling as an alternative transportation method--and the manufacturers as to whether bicycles were a hazard or whether the index merely reflected the widespread usage of a product. If hazards could be identified in the product or its use, both groups wished to eliminate or reduce them.

This report presents the results of a study designed to determine the adequacy of the hazard index as a measure of risk of a product. Section II summarizes the conclusions and recommendations. Section III presents an evaluation of the data collection system from a statistical viewpoint, dealing with the sampling plan and methodology, the data management, and the subsequent in-depth investigations of accidental injuries resulting from products identified for further study. Section IV deals with the uses and misuses which have been made of the NEISS data. It points out that the severity scales and age weightings used in calculation of the "Hazard Index" are subjective, and that the ordering of products is sensitive to which system of weights is used. A number of alternative weighting schemes are suggested, and it is shown how these can be determined empirically as well as subjectively. Section V relates the general conclusions to bicycles specifically, addressing the causality question and other product involvement. Section VI presents a suggested standard analysis or summary of NEISS data for injury accidents associated with bicycles for the years 1972 to 1974. Other relevant literature on bicycle safety is summarized in Section VII, and conclusions and recommendations are presented in Section VIII.

The basic conclusions of the study are as follows:

(1) The hazard index is currently quite dependent on exposure and the particular severity weighting, and interpretation of it as a measure of risk associated with a given product is incorrect.

(2) Injuries associated with bicycles, after adjustment for seasonal effects, have shown only a very slight increase over the three years, much less than the presumed increase in exposure.

It is hoped that the suggestions presented will prove useful to both manufacturers and the CPSC by increasing the utility of data collected by the NEISS.

II. SUMMARY OF CONCLUSIONS AND RECOMMENDATIONS

(1) The Consumer Product Safety Commission has no control over the primary data collection, and reports are based on existing emergency room records as generated by the hospitals. We recommend that as soon as possible the primary data collection be made a responsibility of personnel trained and supported by the CPSC.

(2) The NEISS sample design is complex and difficult to comprehend. It is, therefore, unlikely to gain acceptance or belief by the manufacturing community or among medical research professionals. The CPSC has not determined the possible effects of substitutions of hospitals in the sample on the NEISS estimates. Population shifts which could affect the accuracy of the national estimates have occurred. This is still a minor problem but one which will grow. In view of these conclusions and in light of the difficulties and expense involved in designing and constructing the sample, we recommend that a new sample be designed and planned for implementation as soon as possible after the population data from the 1980 census can be incorporated.

(3) The variance estimates as given by the formulas provided by the CPSC are incorrect. There is a considerable question as to whether accurate estimates of sampling

error are possible in view of the substitution of hospitals and the fact that only one stratum was selected from each superstratum at the first stage. We recommend that careful consideration be given to the design of the new sample to ensure that estimates of the variability are available and that such estimates be reported whenever estimates based on the NEISS data are reported.

(4) The injury categories used are open to question from both the severity scaling point of view and the consistency of application. We recommend that the CPSC consider coordinating their injury categories with those developed by the American Medical Association in the Abbreviated Injury Scale.

(5) The numerical scale used for defining injury severity is a subjective and arbitrary one. Changes in this scale result in changes in the rank orderings of the products. By the nature of the scale it is only useful for ranking the products. Since it is subjective in nature, we recommend that a major effort be implemented to construct a scale acceptable to consumers, manufacturers, and physicians. This will require cooperative effort from all three groups.

(6) The age-weighting used in the construction of the hazard index is subjective and open to question. Any age-adjustment should take into account both the susceptibility by age to accident and to injury once an accident

has occurred. A single age-adjustment over all products appears to be too simplistic to be very useful. We recommend that ranking or comparison of products be done within age groups. If an age adjustment is to be used, we recommend that it take into account susceptibility both to accident and injury and that it be based on available empirical evidence rather than purely subjective judgment.

(7) The NEISS data are obtained from hospital emergency rooms only. Many accidental injuries are treated in physician's offices or directly admitted to a special unit of a hospital. The system should be expanded to include data from these sources since choice of treatment facility is related to location and socioeconomic status.

(8) The Consumer Product Hazard Index has been widely publicized and misinterpreted by the media. To attempt to remedy this, the CPSC should accompany each Hazard Index with a publication of the results of the in-depth investigations which were conducted as a result of a product's ranking high on the index. The conclusions, actions, or need for further information as measured by the assignment of effort for further in-depth study of the product should be presented.

(9) Reports in the past have been based only on the number of cases actually reported through the NEISS. This ignores the probability sampling of the hospitals. Reports should be based on estimates of nationwide injury totals.

(10) The hazard index has been changed several times. These changes have been improvements, and additional changes will be made. When a change is made, the new method should be used to calculate the index for previous years, so that trends over time can be observed.

(11) Based on comparable values, the ten products investigated in this study show a nearly constant trend over the three years considered.

(12) As currently constituted, the hazard index seems more to reflect the exposure or use of a product than the risk associated with that product. We recommend that a measure of hazard be developed which includes a component related to product-specific exposure.

(13) National estimates of monthly totals of bicycle-associated accidents show a very pronounced seasonal component, indicating a large dependence on exposure. Time series models incorporating seasonality, trend, and regular effects should be developed for each product for use in evaluating the effect of any Commission actions. Such evaluation should be planned before any action is implemented.

(14) Bicycle-associated accidents show only a small increasing trend over the three years--less than the presumed increase in exposure would predict. This points out again the need for the development of exposure data.

III. EVALUATION OF NEISS DATA COLLECTION SYSTEM

A. The Sampling Plan

The demographic data upon which the sampling plan is based were drawn from the 1960 U.S. Census. In view of the time during which the planning was done, this is quite understandable. However, there was a 13.3% increase in population in the U.S. from the 1960 to 1970 census. Further, and perhaps more important, there were geographical shifts of the population and also shifts from rural and central city to suburban dwelling. Since the place where a person lives will influence to some extent the products he uses, these differences may be important. An additional effect on the NEISS data is that the weights for the strata are no longer correct. This means that estimates formed from the data will be biased. The direction and magnitude of this bias is unknown; it is probably not a major problem. However, for a system which emphasizes "real-time" data, more up-to-date census information should be incorporated eventually. Some compromise between updating and maintaining the system needs to be reached. The sample cannot be changed too frequently due to the time and effort involved in such a change. Neither can it remain the same indefinitely, however, since population shifts will eventually affect the product-associated injuries to the extent that any inferences drawn would no longer be defensible. The use of the 1960 census in the original plan does not appear to be a major drawback, but it does introduce some bias into the estimates. We would suggest that a new sample be drawn as soon as the 1980 census figures are available.

The use of hospital emergency rooms as the data collection sources has some obvious drawbacks: not all persons injured are treated there, and those who are are likely to differ from those who are not. CPSC estimates that about 40% of injuries are treated in ER, the others being treated in physicians' offices, at home, or elsewhere. To the extent that the socio-economic variables of the group treated at emergency rooms differ from the others, and to the extent that these differences influence these persons' uses of products, the data obtained from the NEISS will be in error. On the other hand, ER's represent a good starting point for the collection of such data. However, these data cannot be blindly accepted as representative.

As a starting point, sampling of the hospital ER records was logical. However, it leaves the system vulnerable at the most crucial stage of any data operation--the control of the basic data. An emergency room is not the best place to collect data in any event. This is particularly true in the case of a serious or life-threatening injury. The ER personnel are correctly much more concerned with the immediate treatment of the injury to assure the survival of the person than they are with the collection of data on how the injury occurred or what product may have been associated with the injury. In general, one would expect more complete and more accurate data for the more minor injuries, since the ER personnel would have more time to record the data.

The staffing of each hospital's emergency room and the training and interests of the staff could easily affect the data collection. At present the NEISS is a sample of hospital emergency room records. (The sample is of the hospitals--within each hospital a census of the ER records is compiled.) This leaves the data subject to variability among hospitals in terms of the proportion of ER records which record product association, as well as other variables. It is also possible that the staff of some hospitals may be particularly sensitive to injuries associated with particular products. This may occur from personal experiences or simply from differences in familiarity with different products. A major improvement in the NEISS would be for the CPSC to train the person or persons responsible for the primary data collection from each ER visit. This would ensure that coverage of product association is the same at all hospitals. It would also increase the acceptance of the results based on the NEISS data. While we appreciate the problems inherent in primary data collection in the setting of an emergency room, the control of the primary data seems a key issue in terms of its quality. At present, the NEISS data represent a sample of existing records. As such they are subject to all of the faults of existing records, the most crucial of which is to what extent the existing records as collected represent the desired population.

The design of the sampling plan seems to be unnecessarily complex. Appendix A gives the detailed description of the sampling plan and procedure used by the CPSC in developing the NEISS. We would have suggested stratifying the U.S. by geographical location to ensure coverage of the entire U.S. (since product use and hence injuries are associated with geographical region and climate, and since the surveillance nature of the data desired would make it desirable to ensure coverage of products with only local distributions) and stratification by size of institution to improve precision. Such a simpler plan would also have advantages in that it would be more readily accepted by the industries concerned. The wider the acceptance of the sampling methodology and resulting data, the more effective the total system could be expected to be in terms of its primary goal of preventing and reducing the severity of accidental injuries.

The plan currently used seems to reflect a confusion between two-stage sampling and stratification. Once strata are defined, a probability sample is taken from each stratum. Each such sample must consist of more than one unit in order to allow for variance computations. In two-stage sampling, the sample at each stage must allow generalization to the population from which the sample was taken. Further, each stage of sampling contributes a term to the variance estimate. Consequently at each stage, when a sample is chosen, it must consist of at least two units from whatever population the units at that stage are drawn from in order that variance estimates may be formed. This has not been done with the current sample. At the first stage, only one "stratum" was selected from each "superstratum." If the superstratum contained only one stratum, this is acceptable, but when the superstratum contained several strata, this seems to preclude estimation of that component of the variance.

A further feature of the current plan is that at the first stage, strata, which were geographical in nature, were selected with probabilities proportional to population. At the second stage, institutions were sampled with probabilities proportional to the number of ER visits. Since the number of ER visits would be highly correlated with the size of the institution and with the population of the surrounding area, the population seems to have been taken into account twice. While sampling on a population proportional basis is well accepted and is a valid method of ensuring that the sample concentrates on high population density areas, where presumably the larger numbers of accidents occur, this may have resulted in more than intended emphasis on high population areas.

Within the 13 geographical strata formed from the 2,300 institutions remaining after the first stage selection, institutions were stratified according to size as defined by number of ER visits. This stratification, however, was peculiar to each geographical stratum, since the largest-institution strata consisted of those institutions which, taken together, comprised 50% of the ER visits in that stratum. Thus the actual sizes of the institutions used in the different geographical regions was different. This is probably not a major problem, but it does make interpretation of the institution size as a stratifying variable rather difficult.

The actual probabilities with which each institution (of the 2,300 left in the subset of the U.S. to be sampled) was selected are extremely complicated and can only be computed with a great deal of effort and computer time. One severe drawback of this is that revision of these probabilities becomes very difficult. In the design of the sampling plan no consideration was given to the practical problem that not all of the institutions selected might agree to participate, or that some of the institutions might subsequently drop out. In actual practice, only about 60% of the hospitals selected by the sample agreed to participate. The other 40% represent second, third, or further choice replacements. Further, a number of insititutions have dropped out, reducing the proportion of the original sample. The result of the nonparticipation problem is that the weights calculated for the institutions in the sample are not correct. This does not seem to have been realized by the CPSC. Further, there seems no way of calculating the correct weights. As a result, all of the projections to national estimates are in error. The magnitude of the errors is unknown, although they are not likely to be extremely large. However, basing policy decisions on data which are known to be erroneous, without a solid estimate of the amount or direction of error, is certainly undesirable and should be corrected.

B. Data Management

Even if the sampling plan used by NEISS were sound, the collection and management of the data leaves much to be desired. The basic data are obtained from the individual ER records. Although this does entail the use of several different forms, this is not a serious problem so long as all of the forms contain the necessary information and so long as all are accurately recorded. However, the person who is responsible for obtaining the data is usually the admitting clerk in the ER. This is a low-skill position with a large turnover rate. Further, these persons have had little if any instruction from NEISS. Yet all the data to be obtained depend upon them. The data that they record on the ER form are the only data available to code and send to the NEISS computer.

ER personnel in general tend to be primarily interested in treating the patients, particularly if the injury is at all serious. Consequently, it is easy to imagine that if there is any rush at all, the determination of product information--which does not appear to be immediately related to the treatment--may be slighted. To the extent that this occurs, data are of poor, perhaps unusable, quality.

It is also possible for various biases to arise through this method of data generation. The CPSC has over 900 categories of products. Few persons could remember all of these with equal facility. The products which are most familiar will get the most attention in determining whether an injury was product-associated; also, any products identified by NEISS for special consideration will likely turn up more frequently than they would had they not been singled out for attention.

A rather different set of data would be generated as a result of asking in the emergency room, "How did this injury occur?" or asking "Was there a product involved in this injury?" than would be obtained from a series of questions: "Was a refrigerator involved in this injury? Was a bicycle involved in this injury?, Was a flight of stairs involved in this injury?,..." which would cover all 900-odd products of interest. If different methods of asking the questions were used at different hospitals, there would be some differences observed. More importantly, the question of how the product identification should be obtained remains.

Assuming that there are no conscious biases on the part of those who fill out the ER records, there still remains the strong possibility of unconscious bias to record some products more frequently than others. At the very best, the basic, primary data collection seems inadequately controlled and can be expected to result in different practices at the different institutions.

Once the ER records have been completed, once a day, a person--different from the one who filled out the records initially--codes them and prepares them for transmission to the computer in Bethesda. The position of the person who does the coding varies from institution to institution, from a junior clerk to a hospital administrator. The quality of the preparation is not in question. However, in the case of severe, life-threatening injuries, the person doing the coding may recall the outcome of a case whose status has changed from the ER. For example, a severely injured person may have died, or some one who was treated and released may have returned and been admitted. If such knowledge is used in the coding, the data on injury severity will be changed.

The interactive data management via computer messages appears to be quite good. However, this merely prevents inaccuracies from entering the data during the coding or computerizing stages. It cannot improve data which were wrong or missing to begin with.

The set of data elements which is collected is quite short. Consequently it does not include as much information as would be desirable. One such major gap is in the details of the injury. Injury is coded only in the NEISS matrix (see Appendix B). Further, only such injury information as is available in the ER is usually included. Later diagnoses or developments are not included in the record unless by chance the coder happens to know of that particular case and changes the ER record. This makes it impossible to determine how a product may have caused an injury, or which injuries were associated with which products in the case of multiple product involvement.

The most crucial gap in the information, however, is the complete absence of any information on how the product was related to the injury. If such injury information is to be used to determine the relative safety of a product, then, logically, the injuries to look at are those which are in some way product-caused. Injuries caused by defects in design or construction would clearly call for remedial action on the part of the manufacturer. Injuries caused by misuse, or improper use, or inadequate maintenance might have been prevented by safety information warning against the hazards of such misuse or mis-maintenance. There is a need for data which would relate an injury to the mechanism which caused it, so that preventive measures might be devised.

A final crucial element missing from the data set is information on exposure. The more persons who use a product, and the more that they use it, the more persons who will have accidents while using that product. Thus, the more popular and useful a product, the higher its raw frequency of "associated" injuries will be. An investigation would probably show that 80% or more of all injuries are "associated" with the underwear of the injured person, since probably at least this high a proportion of individuals were wearing underwear at the time of their accident. In order for meaningful comparisons to be made, a measure of risk incorporating exposure must be used in addition to the frequencies and severities of injuries associated with products. Further, the association must not simply be a trivial one.

One other potential bias in the data collection methods needs to be pointed out. For a given injury, young children are more likely to be brought to an ER for treatment than are older children or adults. Further, a parent or guardian bringing a small child in for treatment is more likely to give adequate information on the toy or product associated with an injury than is an injured person himself. Consequently, unless some measures have been taken to correct for or remove these biases, one would tend to find young children over-represented in the ER population. Also, if this is true, one could expect the product association to be more reliable for the younger children. One would tend to see fewer product associations among older persons. One would also see a more severe distribution of injuries among older persons.

C. In-Depth Studies

No decisions as to regulatory actions the commission is empowered to take would be justifiable on the basis of NEISS data alone. NEISS has the function of sensitizing the commission staff to problems which may exist. Further steps taken before the institution of any regulatory activity include: (1) product testing in either their own or independently operated laboratories; (2) consultation and discussion with the product manufacturers, and (3) more detailed studies of specific cases identified by NEISS as potentially important.

The last of these, the so-called "In-Depth" studies, since they are a subset of the cases occurring in the NEISS data set, are restricted in their efficiency by the limitations of NEISS itself. They are considered necessary, however, because only by such studies can the actual causes of accidents can be determined. Among the causes of product-associated accidents are defects in design or manufacture, inadequate or misleading information on product use, improper or non-intended use of the product, or merely the proximity of the product at the time of the accident. Each of these possibilities, or a combination of them, implies a different type of action to be considered on the part of the commission, and this information as to causality is vital to their deliberations.

Any evidence acquired through in-depth studies of NEISS cases, while useful for producing insights into possible problems, must at this time be viewed as rather weak from the statistical point of view. The reason for this is, first of all, the sampling deficiencies of the NEISS system itself. Beyond that, however, there appears to be no systematic procedure defined for the selection of cases to be studied in-depth, and the studies are restricted to those where the voluntary cooperation of the accident victims is readily available. Time and budget constraints make it difficult for CPSC to do otherwise at this time, but if this class of activity is to attain a status other than as a stimulator to intuition, considerable thought must be given to its proper design and execution.

The Commission has recently instituted a training program for the field investigators who will be engaged in this phase of their work. This will add a great deal to maintaining consistency of data quality among locations. With some assurance in this dimension of the problem, the next step would be to institute a valid sampling plan within the in-depth investigation process.[1] This would allow some statistical conclusions to be drawn about the causative relationships between product design and use and the accidents associated with them.

[1] Since the original draft of this report, the Commission has awarded a contract for the design of the in-depth investigations and the design of the method of selection of cases for in-depth investigations. The study plan calls for demonstration of the procedures by application to three products--power mowers, bath tubs, and architectural glass. The results of the study should lead to a major improvement of this area of the Commission's activities.

IV. USE OF NEISS DATA

Two principal uses for the data generated by the NEISS have been envisioned. The first is a real-time epidemiological surveillance to identify hazards as they occur, and the second is a decision-making tool to identify products for more in-depth studies to lead to a standard or other preventive procedure. A third, somewhat incidental use of the data is to provide a national estimate of the emergency room accident population for in-scope products.[1]

The surveillance use of the NEISS data is somewhat dubious. It must be remembered that for surveillance, the NEISS data are to be supplemented by consumer reported product injuries or product defects, as well as manufacturer's reports of defects, etc. However, with only 120 participating hospitals it seems rather unlikely that the system would be capable of quickly identifying a newly introduced product with a definite hazard. For example, a defective batch of electric toasters could easily all be shipped to one particular geographical region, which might not include a hospital in the NEISS. Even if the hazardous product was introduced in an area covered, it seems unlikely that there would be enough injuries to bring it to attention. In fact, dramatic differences in reported injuries for a

---

[1]"In-scope" products refers to consumer products for which the CPSC has regulatory jurisdiction.

given product (e.g., gas space heaters)* have been observed among reporting hospitals, but this has not been treated as a "product-associated epidemic." It did, however, lead to collection of in-depth reports on the products. The fact that the NEISS deals only with existing ER records seriously limits its rapid response capability. Such crucial information as year of manufacture, causation of injury, and model or brand of product is not available, which would seem to preclude identification of a new hazardous product.

In terms of estimating and describing a national accident population, excluding automobile and industrial injuries, the NEISS is the best source of data extant in the United States today for emergency room visits. The data are still limited to existing records and subject to their incompleteness, non-coverage, and variability by institution. The sample design was such as to make variability estimates very tenuous. These two areas--control of basic data collection and improvement of capability to estimate variances--represent the areas for greatest improvement in the NEISS. Hopefully, both would be much improved if a new sample is selected when the population figures from the 1980 census become available. Nevertheless, the national estimates based on the NEISS are the

*GAMA Analysis of CPSC Data on Gas Space Heaters, Gas Applicance Manufacturers Association, February 14, 1965.

best currently available and represent a useful resource which should find increasing utilization as more researchers become aware of their availability.

The major use which has actually been made of the data has been to order products for attention by the CPSC. This has resulted in the "Consumer Product Hazard Index" published at various times. The intent of this index was to aid the CPSC in determining where to allocate effort for the more detailed epidemiological (in-depth) investigations to determine whether an unreasonable hazard existed, and, if so of what nature. Based on the results of this investigation, appropriate remedial action would be suggested. Such action could be an outright ban for a dangerous product, a safety design standard to improve the safety and utility of a product, an educational campaign to reduce misuses of the product, etc.

Unfortunately, the "hazard index" has been widely publicized by various media with the resultant--though unjustified--conclusion that those products ranking high on the list are somehow dangerous or unreasonably hazardous to the users. Adding to the confusion is the fact that the calculation of the index has involved at least four major changes, each of which makes the index noncomparable at different times. The changes have generally made products appear to be becoming more hazardous with time. However, this impression is an artifact of the different hazard indices. Based on comparable values, most products (of ten considered) actually have quite flat profiles. That is, no apparent changes have occurred except for year-to-year variation.

Since the hazard index has been widely publicized and no doubt will continue to be widely publicized, it is important to understand its calculation. This has changed at different times and will probably continue to change. Briefly, injuries are classified according to a subjective severity scale into seven different classes. A severity score for each class has been subjectively determined by the CPSC. These are:

Severity class	1	2	3	4	5	6	7
Severity score	10	12	17	31	81	340	2516

Originally the hazard index was to be calculated by multiplying each of these scores by the observed number of injuries in each class, and doubling if the injury occurred to a person under ten or over 65. Subsequently the doubling for the over-65 group was eliminated. (The doubling for young ages is referred to by the CPSC as an "age adjustment.") Following this method of calculation, the index was reported separately by the Commerce Clearing House (not the CPSC) for: (1) children under five, (2) children under 16, (3) women over 16, and (4) men over 16. Products were ranked separately in the four groups.

Recently other changes have been made. The age adjustment is now to multiply by 2.5 for injuries occurring to children under 15 years of age. Further, the number of injuries in each severity category is no longer the observed number of injuries, but is now the estimated national total of such injuries. Consequently, current numerical values of the hazard index differ by several orders of magnitude from earlier ones.

A. Frequency-Severity Estimates

CPSC deemed it necessary, for purposes of cross-product comparisons, to have a single number represent the degree to which a specific product was responsible for the cause of injury in the population. The result of this assumption is the production of a frequency-severity index--that is, the sum for each product of the number of injuries, each multiplied by its severity. Expressed in mathematical terms:

$$I_{FS} = \sum_{i}^{n} S_i$$

where I_{FS} is the frequency-severity index for a specific product. S_i is the severity attached to the i^{th} injury, and n is the number of reported injuries for the product.

While having a considerable amount of face validity, such an index has obvious limitations. First, the frequency involved was merely the number of observed injuries. This ignores the sampling scheme by giving equal weights to all hospitals. Secondly, if as is currently the case, the frequencies are used to generate an estimate of the nationwide index for the occurrence of such injuries, the estimate is subject to the effectiveness of the sampling plan in achieving representativeness. This has been discussed above. Thirdly, as will be discussed below, the severity index itself is a subjective, arbitrary one, and there is no apparent justification in combining these injury data in the form of a linear additive.model. The presumption, for example, that an injury at level 4 is three times as bad as an injury at level 1 cannot be accepted without some rationale in terms of such variables as cost, discomfort, disability over defined time periods, days of lost work or lessened efficiency, and the like. Toward the higher end of the injury spectrum, the linear additive assumption in combining the data elements appears even more tenuous.

Several factors need to be incorporated into the index. The frequency of the injury needs to be considered since more frequent injuries are of greater interest for preventive measures. The severity of the injury must also be considered, since it is more important to prevent life-threatening injuries or disabling injuries than minor injuries. Injuries which occur predominantly to special groups (the very young, the very old, predominantly to women, etc.) might be of special interest. Finally exposure should be considered. An injury which occurs once every 1,000,000 uses of a product, on the average, should be less important than one which occurs once every 100 uses of the product, on the average, even though the former might result in more total injuries annually due to much greater use of one product than another.

B. Age "Adjustment"

The CPSC, feeling, and probably correctly so, that injuries to different age groups were of different importance, thought it wise that the indices produced by NEISS be adjusted or weighted for age. The original decision was to give a weighting factor of two to injuries occurring to individuals ten years old or younger and to those sixty-five years of age and older. The weighting has subsequently been removed from the data pertaining to the elder of these two groups.

This weighting scheme is subjective; no rationale is given for using it other than some obvious common-sense criteria. As to the younger group, it is assumed that CPSC is concerned, as is our society in general, with adequately protecting children. The weight then is present to help identification of products which are more dangerous to children. It is well known, however, that the younger group will suffer less injury and recover faster from the same trauma than will an older group, and it is nowhere evident that a weight based on age alone, rather than a combination of age and type of injury, will do anything toward making a useful adjustment. In the case of the older victims, the reverse is true. The same trauma may well produce injuries which are both more severe and more costly to the individual and to society, and yet the attempt to get at this phenomenon by an increased weight was discarded.

Currently the "age adjustment" is to multiply the severity score by 2.5 if the injury is to an individual under 15. The rationale for the number 2.5 (more accurately 2.52) is that there are approximately 2.5 times as many persons in the U.S. who are older than 15 as who are under 15. Thus the factor of 2.5 is intended to make the totality of accidents to persons under 15 comparable to the totality of accidents to persons over 15. Implicit in this is a value judgment that injuries to persons under 15 are more important than are injuries to persons over 15. Alternately it may reflect the judgment that injuries to adults are not very important. This remains a subjective judgment by the CPSC that it is more important to protect children and that the weighting of 2.5 for persons under 15 will adequately accomplish this. It is our opinion that such an age adjustment is overly simplistic; we believe that it is more useful to consider various age groups separately, because they tend to use different products.

The use of the term "age adjustment" may mislead some persons. This term is generally used in comparisons of groups to ensure that observed differences are not due simply to different age distributions in the two groups. In this context it is also referred to as a standardization for age or as the calculation of an age-standardized rate. When used to make the morbidity or mortality experience of two or more populations comparable (with respect to age), an age adjustment proceeds as follows. The age-specific rates are estimated for each of several age groups in each population. These age-specific rates are then applied to a standard population to determine how the rates would compare if all populations had the same age distribution as the standard. The selection of the standard distribution can affect the comparisons, particularly if the age-specific rates are much different. However, if the standard age distribution is meaningful (e.g, that of the U.S.), comparisons of the adjusted rates may be more meaningful than comparisons of the original crude rates.

No such adjustment is meant by the "age-adjustment" in the hazard index--the age-adjusted frequency severity index (AFSI). Instead, it might more properly be referred to as an age weighting. In effect, different injury severity scores are used for different age groups because of the subjective judgment that the different age groups are of varying importance. This judgment of the differing importance of the age groups is open to question. The different weights applied to the age groups are even more open to question. Since different choices of the weights result in different product orders, a particular selection must invariably--consciously or unconsciously--result in an unfavorable comparison for those products which have higher age distributions in the age groups which get the higher weights.

If an age weighting or adjustment is to be applied to all products uniformly, it would seem logical for this to reflect either a concern for protecting individuals more likely to be seriously injured by an accident or protecting individuals who are more likely to become involved in an accident, or both. Whichever reason is used, it is not necessary to base the adjustment purely on a subjective determination. If over-involvement is the rationale, the appropriate adjustment can be estimated from data on age distributions of accident victims compared to age distributions for the population as a whole. If susceptibility to injury is a criterion, an age adjustment can be based on age-specific tolerances to a specified injury or to a cruve of the same general shape as a specific mortality curve. If both are to be considered, an additive or multiplicative combination of them could be considered.

It is possible, of course, that the various possibilities for age adjustment might make little or no difference in the relative ranking of the products. If that were the case it would not matter which was used or if any were used at all. However, it turns out that the ranking of the products does depend heavily on which sort of adjustment is used. Consequently the adjustment should be based on as widely acceptable a rationale as possible. This implies that some degree of objectivity should be incorporated.

C. Scheme for Weighting Severity

The operators of NEISS are also the prime users of the system, and one of their chief objectives is to detect problems with specific products. The frequency of associations of products with injuries is not considered alone to be a good enough indicator of a problem area. The severity of injury would affect their consideration, and since more severe injuries would indicate more severe problems, it was decided to weight the severe injuries more heavily than less severe ones. The weighting scheme used is subjective, and the NEISS people are quoted as saying that it "has an algebraic rationale but not a real-life rational." It seems unlikely that an arbitrary weighting scheme can be said to have an "algebraic rationale," but what this is intended to imply is that the formula for calculating the weights is constant. After arbitrarily assigning a weight of ten to a severity designation of 1, the other weights are calculated by:

$$W_i = W_{i-1} + (.10)(2^{i-1})(W_{i-1})$$

where W_i is the weight of the i^{th} severity index. This formula produces the following table:

i	W_i
1	10
2	12
3	17
4	31
5	81
6	340
7	2516
8	34721

The last entry in the table has recently been revised, and the weight for the highest severity index has been assigned the value of 2516.

Again, the particular severity scale which is used may not affect the hazard index. If the relative ranking of products is not sensitive to which scale is used, it would make little difference which was used. However, we have found that different severity scores affect the rankings of the products. Again, this means that the index is sensitive to which particular scale is used. Consequently, the scale should be carefully developed and defensible.

Tables 1-14 present the rankings of ten selected products based on several different severity scales and age adjustments, for three calendar years, 1972, 1973, and 1974. Within a particular scale, rankings vary somewhat by year. Also presented is a three year profile for each product illustrating the trend in that injury scale for that product. In addition to the various severity weightings used, several age adjustments have been considered. One of these is an involvement measure developed by forming the ratio of the percent injured by age (in a set of severe burn injuries) to the percent of the population in that age group. Two measures of susceptibility to injury were used. One is based on the tolerance to burn injuries[1] by age, and the other is based on the age-specific mortality rates for the U.S.[2] Graphs of the various age adjustment curves are presented in Figure 1-3.

Tables 1-14 give the resulting rankings of the ten product groups (described in Table 15) for the three calendar years. The age weightings used are illustrated in Figures 1-3. For some of the tables (9, 11, and 14) no different weightings for different age groups were used.

---

[1]Feller, I., Flora, J., and Bawol, R. (1975). Baseline Survival Curves for Burn Injuries, submitted to the Journal of the American Medical Association.

[2]Shyrock and Siegel, "The Materials and Methods of Demography", Bureau of the Census (1973).

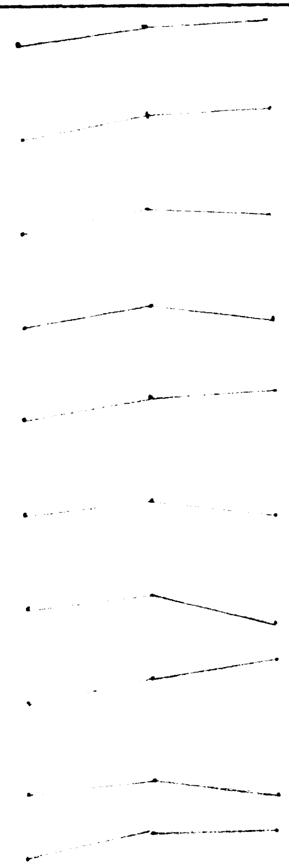
TABLE 1. NEISS INJURY SCORE - AGE-ADJUSTED FOR SUSCEPTIBILITY TO TRAUMA

	Estimated National Projections						Observed Injuries					
	Product Rank			Product Trend			Product Rank			Product Trend		
	1972	1973	1974	1972	1973	1974	1972	1973	1974	1972	1973	1974
Bicycles	1	1	2				2	2	2			
Stairs	2	2	1				1	1	1			
Doors	4	3	4				3	3	3			
Swings, Slides	8	7	6				7	7	6			
Liquid Fuels	5	6	5				9	5	5			
Architectural Glass	6	8	8				6	8	7			
Power Lawnmowers	3	4	7				4	4	8			
Baseball Equipment	7	5	3				5	6	4			
Baths	9	9	9				8	9	9			
Skates, Skateboards, Scooters	10	10	10				10	10	10			

TABLE 2. NEISS INJURY SCORE - AGE-ADJUSTED FOR PROBABILITY OF INVOLVEMENT

	Estimated National Projections Product Rank			Observed Injuries Product Rank		
	1972	1973	1974	1972	1973	1974
Bicycles	1	1	2	2	2	2
Stairs	2	2	1	1	1	1
Doors	3	3	3	3	3	3
Swings, Slides	6	4	5	4	4	5
Liquid Fuels	4	5	4	7	5	4
Architectural Glass	7	7	7	6	6	6
Power Lawnmowers	8	8	8	5	9	9
Baseball Equipment	9	6	6	9	8	7
Baths	5	9	9	8	7	8
Skates, Skateboards, Scooters	10	10	10	10	10	10

Estimated National Projections — Product Trend 1972 1973 1974



Observed Injuries — Product Trend 1972 1973 1974

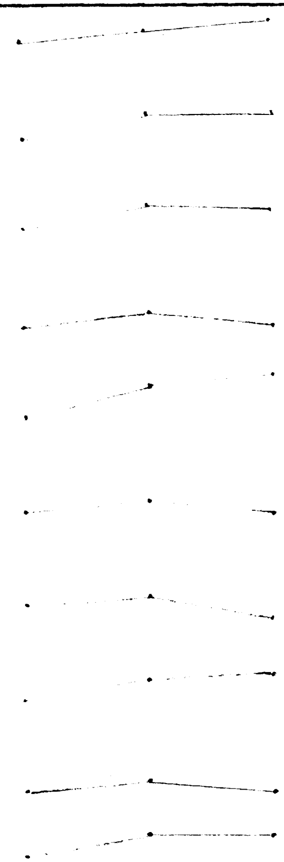


TABLE 3. NEISS INJURY SCORE - AGE-ADJUSTED FOR SUSCEPTIBILITY TO TRAUMA AND PROBABILITY OF INVOLVEMENT

	Estimated National Projections						Observed Injuries					
	Product Rank			Product Trend			Product Rank			Product Trend		
	1972	1973	1974	1972	1973	1974	1972	1973	1974	1972	1973	1974
Bicycles	2	2	2				2	2	2			
Stairs	1	1	1				1	1	1			
Doors	3	3	3				3	3	3			
Swings, Slides	6	4	5				5	4	5			
Liquid Fuels	4	5	4				6	5	4			
Architectural Glass	7	8	7				7	7	6			
Power Lawnmowers	8	7	9				8	8	9			
Baseball Equipment	9	6	6				9	9	7			
Baths	5	9	8				4	6	8			
Skates, Skateboards, Scooters	10	10	10				10	10	10			

TABLE 4. NEISS INJURY SCORE - AGE-ADJUSTED FOR GENERAL SUSCEPTIBILITY

	Estimated National Projections						Observed Injuries					
	Product Rank			Product Trend			Product Rank			Product Trend		
	1972	1973	1974	1972	1973	1974	1972	1973	1974	1972	1973	1974
Bicycles	2	2	2				2	2	2			
Stairs	1	1	1				1	1	1			
Doors	4	4	3				4	3	3			
Swings, Slides	9	7	8				9	9	9			
Liquid Fuels	5	5	6				8	5	5			
Architectural Glass	7	8	9				6	7	6			
Power Lawnmowers	3	3	5				3	4	8			
Baseball Equipment	8	6	4				7	6	4			
Baths	6	9	7				5	8	7			
Skates, Skateboards, Scooters	10	10	10				10	10	10			

TABLE 5. UNIFORM INJURY SCORE - AGE-ADJUSTED FOR SUSCEPTIBILITY TO TRAUMA

	Estimated National Projections						Observed Injuries					
	Product Rank			Product Trend			Product Rank			Product Trend		
	1972	1973	1974	1972	1973	1974	1972	1973	1974	1972	1973	1974
Bicycles	1	2	2				2	2	2			
Stairs	2	1	1				1	1	1			
Doors	5	4	4				3	3	4			
Swings, Slides	6	6	6				6	6	6			
Liquid Fuels	10	10	10				10	10	10			
Architectural Glass	3	5	5				4	5	5			
Power Lawnmowers	7	7	9				7	8	9			
Baseball Equipment	4	3	3				5	4	3			
Baths	8	9	8				8	9	8			
Skates, Skateboards, Scooters	9	8	7				9	7	7			

TABLE 6. UNIFORM INJURY SCORE - AGE-ADJUSTED FOR PROBABILITY OF INVOLVEMENT

	Estimated National Projections						Observed Injuries					
	Product Rank			Product Trend			Product Rank			Product Trend		
	1972	1973	1974	1972	1973	1974	1972	1973	1974	1972	1973	1974
Bicycles	1	2	2				2	2	2			
Stairs	2	1	1				1	1	1			
Doors	3	3	3				3	3	3			
Swings, Slides	5	4	5				5	5	6			
Liquid Fuels	9	10	8				10	9	9			
Architectural Glass	4	5	6				4	4	5			
Power Lawnmowers	8	8	10				8	10	10			
Baseball Equipment	6	6	4				6	6	4			
Baths	7	7	7				7	7	7			
Skates, Skateboards, Scooters	10	9	9				9	8	8			

TABLE 7. UNIFORM INJURY SCORE - AGE-ADJUSTED FOR SUSCEPTIBILITY TO TRAUMA AND PROBABILITY OF INVOLVEMENT

	Estimated National Projections						Observed Injuries					
	Product Rank			Product Trend			Product Rank			Product Trend		
	1972	1973	1974	1972	1973	1974	1972	1973	1974	1972	1973	1974
Bicycles	2	2	2				2	2	2			
Stairs	1	1	1				1	1	1			
Doors	3	3	3				3	3	3			
Swings, Slides	4	4	5				5	4	5			
Liquid Fuels	8	8	8				9	8	8			
Architectural Glass	5	5	6				4	5	6			
Power Lawnmowers	9	9	10				8	10	10			
Baseball Equipment	6	6	4				6	6	4			
Baths	7	7	7				7	7	7			
Skates, Skateboards, Scooters	10	10	9				10	9	9			

TABLE 8. UNIFORM INJURY SCORE - AGE-ADJUSTED FOR GENERAL SUSCEPTIBILITY

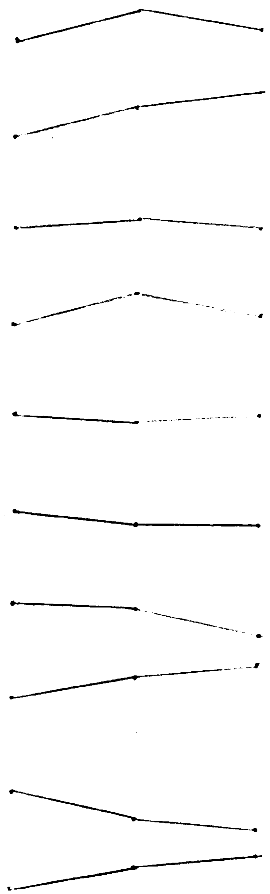
	Estimated National Projections						Observed Injuries					
	Product Rank			Product Trend			Product Rank			Product Trend		
	1972	1973	1974	1972	1973	1974	1972	1973	1974	1972	1973	1974
Bicycles	2	2	2				2	2	2			
Stairs	1	1	1				1	1	1			
Doors	5	4	4				3	3	4			
Swings, Slides	6	6	6				6	6	6			
Liquid Fuels	10	10	10				10	10	10			
Architectural Glass	3	5	5				4	5	5			
Power Lawnmowers	7	7	8				7	8	9			
Baseball Equipment	4	3	3				5	4	3			
Baths	8	8	9				8	7	8			
Skates, Skateboards, Scooters	9	9	7				9	9	7			

TABLE 9. NOMINAL INJURY SCORE

	Estimated National Projections						Observed Injuries					
	Product Rank			Product Trend			Product Rank			Product Trend		
	1972	1973	1974	1972	1973	1974	1972	1973	1974	1972	1973	1974
Bicycles	1	1	1				2	2	2			
Stairs	2	2	2				1	1	1			
Doors	6	5	4				4	3	4			
Swings, Slides	5	4	5				6	6	6			
Liquid Fuels	10	10	9				10	10	9			
Architectural Glass	4	6	6				5	5	5			
Power Lawnmowers	8	8	10				8	8	10			
Baseball Equipment	3	3	3				3	4	3			
Baths	7	7	7				7	7	7			
Skates, Skateboards, Scooters	9	9	8				9	9	8			

TABLE 10. NOMINAL INJURY SCORE - AGE-ADJUSTED FOR SUSCEPTIBILITY TO TRAUMA

	Estimated National Projections						Observed Injuries					
	Product Rank			Product Trend			Product Rank			Product Trend		
	1972	1973	1974	1972	1973	1974	1972	1973	1974	1972	1973	1974
Bicycles	1	1	1				2	2	2			
Stairs	2	2	2				1	1	1			
Doors	6	5	4				3	3	4			
Swings, Slides	4	4	5				6	5	6			
Liquid Fuels	10	10	9				10	10	9			
Architectural Glass	5	6	6				5	6	5			
Power Lawnmowers	8	8	10				8	8	10			
Baseball Equipment	3	3	3				4	4	3			
Baths	7	7	7				7	7	7			
Skates, Skateboards, Scooters	9	9	8				9	9	8			



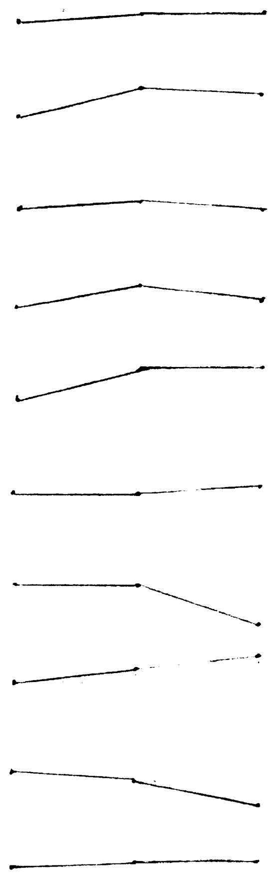


TABLE 11. NOMINAL-SQUARED INJURY SCORE

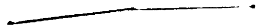
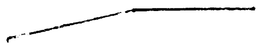
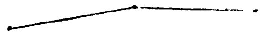
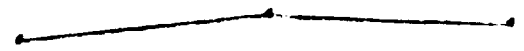
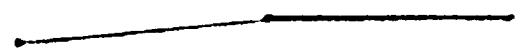
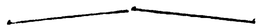
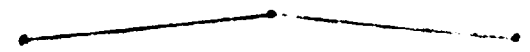
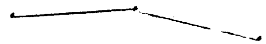
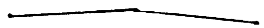
	Estimated National Projections				Observed Injuries			
	Product Rank			Product Trend	Product Rank			Product Trend
	1972	1973	1974	1972 1973 1974	1972	1973	1974	1972 1973 1974
Bicycles	1	1	1		2	2	2	
Stairs	2	2	2		1	1	1	
Doors	5	5	5		5	4	4	
Swings, Slides	4	4	4		4	5	5	
Liquid Fuels	9	8	8		10	9	8	
Architectural Glass	6	6	6		6	6	6	
Power Lawnmowers	8	9	10		8	8	10	
Baseball Equipment	3	3	3		3	3	3	
Baths	7	7	7		7	7	7	
Skates, Skateboards, Scooters	10	10	9		9	10	9	

TABLE 12. NOMINAL-SQUARED INJURY SCORE - AGE-ADJUSTED FOR SUSCEPTIBILITY TO TRAUMA

	Estimated National Projections						Observed Injuries					
	Product Rank			Product Trend			Product Rank			Product Trend		
	1972	1973	1974	1972	1973	1974	1972	1973	1974	1972	1973	1974
Bicycles	1	1	2				2	2	2			
Stairs	2	2	1				1	1	1			
Doors	5	5	5				4	3	3			
Swings, Slides	3	3	4				3	4	5			
Liquid Fuels	9	9	8				10	8	8			
Architectural Glass	6	6	6				6	6	6			
Power Lawnmowers	8	8	10				8	9	10			
Baseball Equipment	4	4	3				5	5	4			
Baths	7	7	7				7	7	7			
Skates, Skateboards, Scooters	10	10	9				9	10	9			

TABLE 13. ABBREVIATED INJURY SCALE-EQUIVALENT INJURY SCORE - AGE-ADJUSTED FOR SUSCEPTIBILITY TO TRAUMA

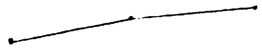
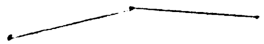
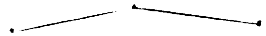
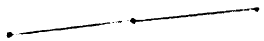
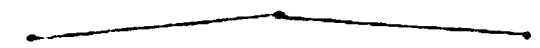
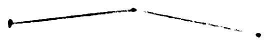
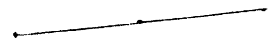
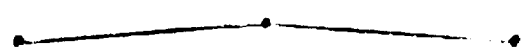
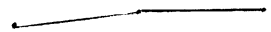
	Estimated National Projections				Observed Injuries			
	Product Rank			Product Trend	Product Rank			Product Trend
	1972	1973	1974	1972 1973 1974	1972	1973	1974	1972 1973 1974
Bicycles	1	1	1		2	2	1	
Stairs	2	2	2		1	1	2	
Doors	5	5	5		5	4	4	
Swings, Slides	4	4	4		4	5	5	
Liquid Fuels	7	6	6		7	7	7	
Architectural Glass	6	7	7		6	6	6	
Power Lawnmowers	10	10	10		10	10	10	
Baseball Equipment	3	3	3		3	3	3	
Baths	8	8	8		8	8	8	
Skates, Skateboards, Scooters	9	9	9		9	9	9	

Table 14 differs from the other tables in that it addresses more clearly the risk of serious injury. As such it needs some additional explanation to be understood. If the combined injury distribution for the ten product groups is taken as a standard or reference distribution, then one may ask whether a bicycle injury is likely to be more severe or less severe than this standard. This question may be addressed by calculating for each product a statistic called the mean ridit.[1,2] The interpretation of this statistic is that it provides an estimate of the probability that a person involved in a bicycle accident will be less seriously injured than one involved in accidents in general. Thus products for which this probability is small tend to result in serious injuries from accidents; products for which this probability is large tend to result in less serious injuries. In contrast to the other hazard indices, this measure is directly interpretable. Consequently, it has been reported in Table 14a. The numbers in this table may be interpreted as the probability that an injury associated with the stated product group is of less than average severity. Here the "average severity" is relative to the combined distribution of severities projected for the combined accidents of the ten product groups.

---

[1] Bross, I.D.J. (1958). "How to Use Ridit Analysis", Biometrics.

[2] Flora, J.D. (1974). "Ridits: A New Look at an Old Technique for the Analysis of Accident Injury Data, HIT LAB Reports 5, No. 3.

TABLE 14. MEAN RIDIT INJURY SCORE

	Estimated National Projections						Observed Injuries					
	Product Rank			Product Trend			Product Rank			Product Trend		
	1972	1973	1974	1972	1973	1974	1972	1973	1974	1972	1973	1974
Bicycles	4	4	4				4	4	4			
Stairs	5	5	6				6	6	6			
Doors	7	7	7				7	7	7			
Swings, Slides	3	3	3				3	3	3			
Liquid Fuels	1	1	1				1	1	1			
Architectural Glass	10	9	8				9	9	9			
Power Lawnmowers	6	6	5				5	5	5			
Baseball Equipment	8	8	9				8	8	8			
Baths	2	2	2				2	2	2			
Skates, Skateboards, Scooters	9	10	10				10	10	10			

TABLE 14A. MEAN RIDIT INJURY SCORE

	Estimated National Projection			Observed Injuries		
	1972	1973	1974	1972	1973	1974
Bicycles	.4896	.4802	.4780	.4800	.4771	.4747
Stairs	.5099	.5128	.5069	.5080	.5178	.5090
Doors	.5349	.5312	.5230	.5200	.5282	.5282
Swings, Slides	.3929	.3922	.3748	.3875	.3798	.3774
Liquid Fuels	.2705	.2978	.3000	.2900	.2788	.3306
Architectural Glass	.5578	.5590	.5401	.5500	.5499	.5545
Power Lawnmowers	.5271	.5287	.4898	.5035	.5056	.5030
Baseball Equipment	.5400	.5431	.5445	.5306	.5442	.5531
Baths	.3224	.3423	.3047	.3165	.3459	.3334
Skates, Skateboards, Scooters	.5534	.5648	.5692	.6504	.5654	.5691

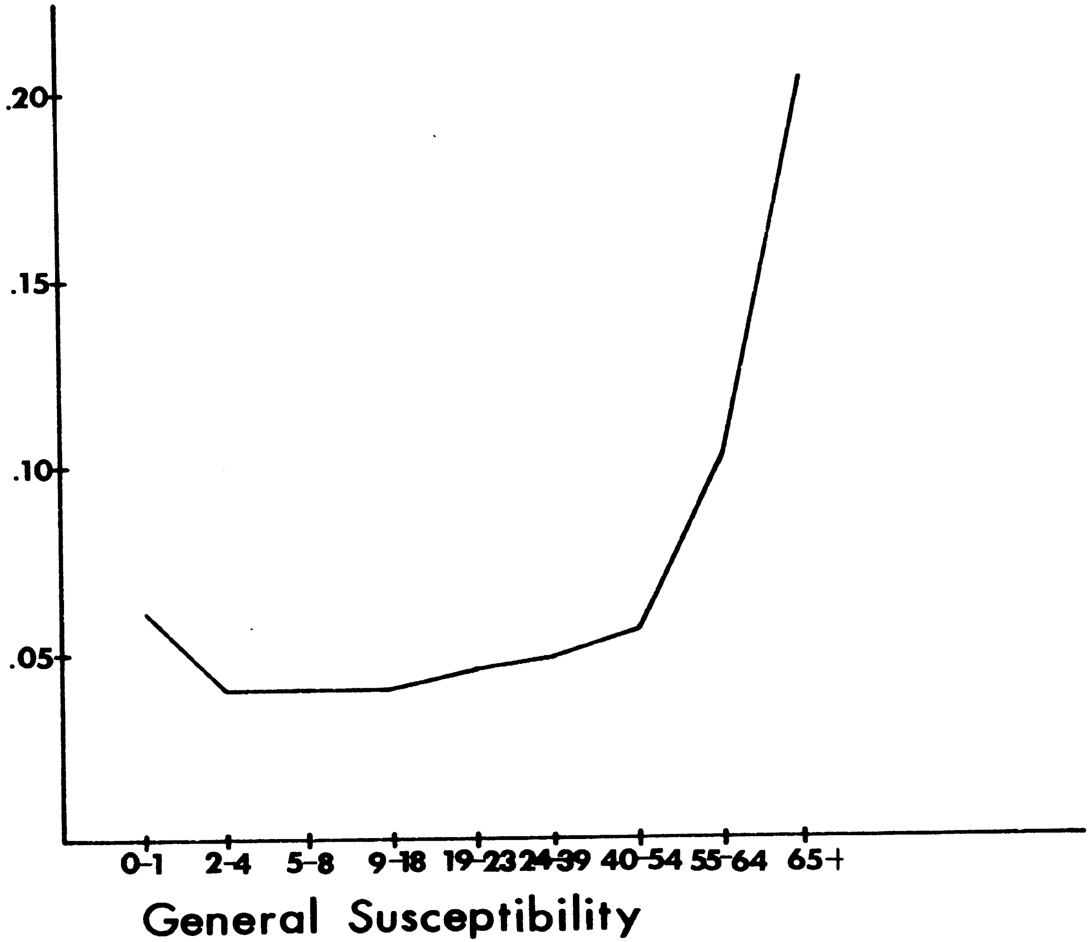


FIGURE 1. AGE WEIGHTING PROPORTIONAL TO AGE-SPECIFIC MORTALITY RATES

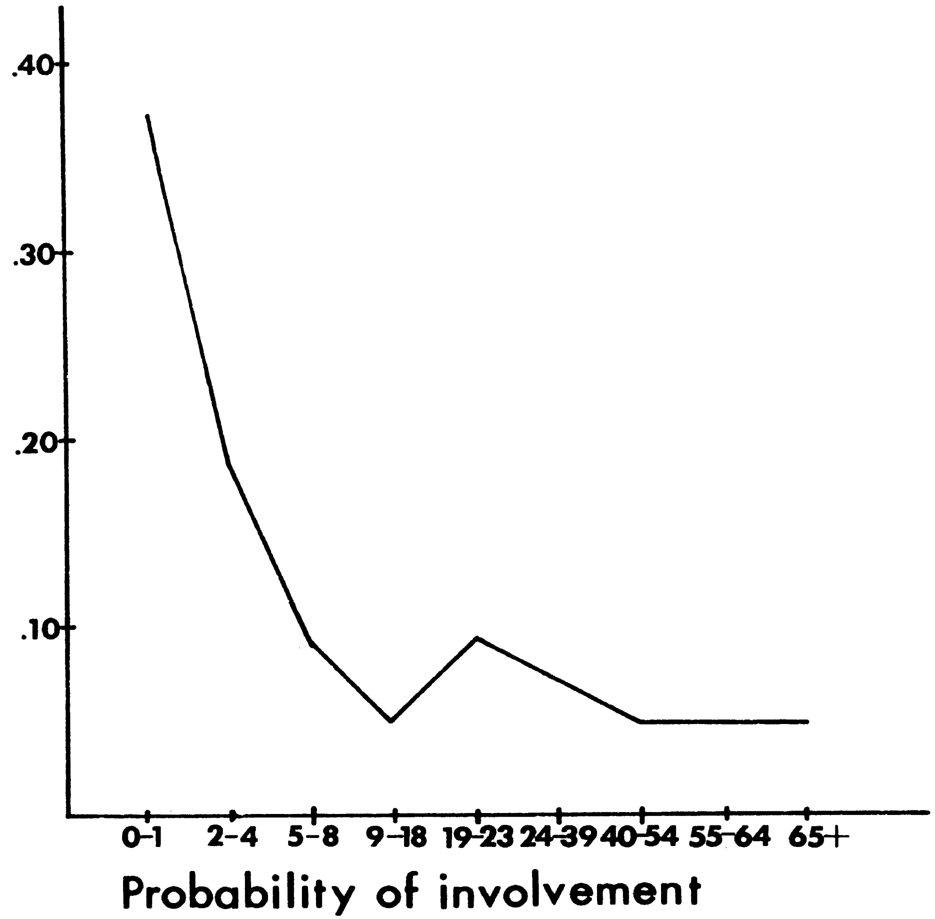


FIGURE 2. AGE WEIGHTING PROPORTIONAL TO LIKELIHOOD OF INVOLVEMENT IN ACCIDENT

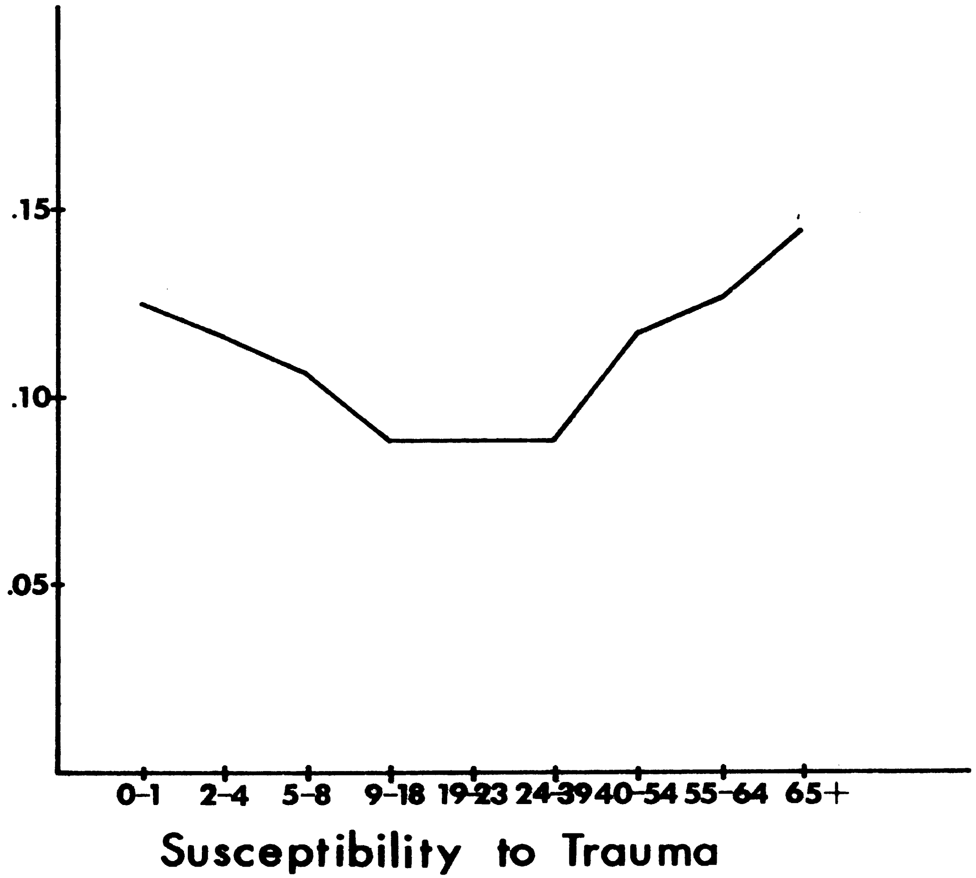


FIGURE 3. AGE WEIGHTING PROPORTIONAL TO SUSCEPTIBILITY TO BURN TRAUMA

TABLE 15. PRODUCTS AND CODES USED IN ANALYSIS

Product Identification	1973 CPSC Rank	Codes Used 1972,1973	Codes Used 1974 on
Bicycles	1	1202	1202
Stairs	2	1801,1840	1840,1842,1843
Doors	3	1805,1822,1827,1858,0117	1822,1827,1844,1845,1846 1847,1848,1849,1858,0117
Swings, Slides	8	1201,1241,1242,1243,1244	1201,1241,1242,1243,1244
Liquid Fuels	9	0910,0911,0940,1247	0910,0911,0940,1247
Architectural Glass	10	1815,1823,1824,1825,1835 1836,1837	1815,1823,1824,1825,1835 1836,1837
Power Lawnmowers	11	1401,1418,1419,1420,1421	1401,1418,1419,1420,1421 1436,1437
Baseball Equipment	12	1204,1341	1204,1341
Baths	14	0611	0611
Skates, Skateboards,	42	1308,1329,1333	1308,1329,1333

Although bicycles rank fourth worst, the associated probability is quite close to one-half, indicating that bicycle injuries are about of median severity relative to this group. This contrasts, for example with the accidents associated with liquid fuels or bath tubs, which are very likely to be more severe. Most of the others, e.g., skates-associated injuries, are not a great deal more likely to be less than of median severity than are those associated with bicycles.

Whatever scheme is used to determine an index, it is clear that it should be based on the national projections of annual injuries. To base it only on the injuries actually reported ignores the fact that a probability sample has been used and treats the hospitals as if they had been initially selected with equal probabilities. It is also certainly the case that the numerical value of the hazard index will be a meaningless number. The measurements are only on an ordinal scale, not an arithmetic scale or a ratio scale. Consequently, the only information to be obtained is the relative ranking of the products and only those ranks--first, second, etc., should be reported.

It is clear that the calculations used to arrive at the ranks must also be explained and the numerical values for each product must be available. However, the numerical values of the hazard index should be rescaled (e.g., by division by a positive constant) so that there is little danger that someone would misinterpret them as the observed number of injuries or the projected number of injuries or given them some interpretation other than simply a numerical score used to rank order products.[1]

The AFSI, whether calculated under the 1973 or the 1975 method, is a subjective scale for injuries resulting from product-associated accidents which led to emergency room visits. As such, the numerical values calculated for products in the scale have no intrinsic meaning. They are not on an arithmetic or a ratio scale, so that it is meaningless to compare two products by subtracting their AFSIs or by taking the ratio of their AFSIs. The only proper use of the numbers is to rank-order the products. This ranking of the products is particular to this subjective scale. We have used other scales for ranking products according to projected injury-severity combinations and observed rather different rankings. For example, the product originally ranked 42nd (in 1973) ranks as high as seventh on some of these scales. Shifts from third to eighth among the ten products are common. The point is that the only use of the AFSI is to rank the products, and that this ranking depends fairly strongly on the particular subjective combination of injury-severity weights and "age adjustment" weights used.

Comparison of the 1973 AFSI to the 1975 AFSI can hardly fail to be misleading since both the weights have changed

[1]While this report was in press, the CPSC agreed to rescale the Hazard Index at the suggestion of the authors and has since done so by dividing the projected AFSI by one million.

and the set of injuries used to calculate the index have changed. Obviously if one forms a weighted sum over 458,000 estimated injuries rather than 20,177 observed injuries, one will arrive at a much larger number. This is the main source of difference between the 1973 and 1975 AFSIs, although the 1975 also uses larger weights.

In view of the potential misinterpretation of the numerical values for the 1973 and 1975 AFSIs, it is our opinion that they should not be published for comparison. This may be equivalent to attempting to prevent a foreseeable misuse of a product--data in this case. We would recommend that the most meaningful information to be published would consist of the following items for each product:

(1) Actual number of product-associated injuries reported through NEISS.

(2) Projected annual number of product-associated injuries for the U.S.

(3) Relative ranking of products on the AFSI scale, but not the numerical values unless these are rescaled to preclude misinterpretation.

(4) Description of how the annual total is estimated, together with a description of the scaling procedure used to rank the products.

(5) Error estimates for the projected annual number of product-associated injuries, together with a statement of what changes in the ranking of products could occur as a result of these errors.

The new method of calculating the AFSI on the basis of the national projections of injuries is clearly preferable to the old method of basing it only on observed injuries. It seemed rather nonsensical to have an elaborate sample design resulting in weights for hospitals ranging from 1.5 to 514.4 and then to perform calculations using equal weights for all hospitals. In that sense the new method is the more sound one. However, it still should have some estimate of error variability reported.

Finally, some measure of exposure should be incorporated into the index. The definition of exposure is a difficult problem. Measuring it may be ever harder. However, some very approximate but still potentially useful measures could be incorporated quite easily. Different measures of exposure might be more appropriate for different products. For example, the number of refrigerators in use is probably the most appropriate measure for refrigerators. On the other hand, exposure for bicycles would be better measured in terms of miles ridden, time spent biking, or number of trips than simply by bicycles in use. Such measures would be very difficult to obtain, of course, but the first approximation of number of products in use should be readily estimable. A measure such as injuries per product in use, or injuries weighted by severity per product in use, might be quite useful.

The CPSC is aware of the need for exposure data. Reportedly they are designing a survey to estimate exposure at least for some of the more important products. This exposure survey is planned to become a part of the National Health Survey conducted by the National Center for Health Statistics. It is hoped that development of the exposure measures and the survey will proceed as rapidly as possible. It seems highly likely that a hazard index which has some sort of a measure of exposure as a denominator will result in a rather different rank ordering of products in terms of risk. It is certain that valid exposure data would enhance the usefulness of the NEISS data.

The development of useful exposure data will be a difficult task. It is not clear at the onset just which measure will be most meaningful. It is likely that different measures may be appropriate for different products. If so, products cannot be directly compared except in terms of some concept of risk. For example, some products such as refrigerators, carpets, etc., may have as the most reasonable exposure measure the number of products in use.

The assumption would be that everyone uses a refrigerator in about the same way and about the same amount. On the other hand, products such as lawn mowers or bicycles might be better related to exposure in terms of hours of use or times used. In the case of bicycles, the use--whether as a toy, riding for recreation, or as a vehicle of transportation--as well as the time, mileage, or number of uses might be important. It is also clear that exposure for many products will be seasonally related. Further, the type of exposure which it proves possible to gather may not be the data which one would desire under ideal conditions. Still, for each product, the exposure measures which would be desirable should be determined and ordered by usefulness. As the survey is designed, and information on which measures can be obtained at a reasonable cost and with acceptable accuracy, a compromise decision can be reached on which exposure measure will be used for each product.

D. Variance Estimation

Whenever an estimate of a population mean or total is made from a sample it is important to provide some estimate of the variability of the estimate or some measure of the reliability of the estimate. In sampling situations this measure is usually the mean square error or the root mean square error. If the estimates are known to be unbiased, this reduces to the sampling variance or to the sampling standard deviation, respectively. This practice should also be followed by the CPSC in estimating the national totals of injuries of various types. To date, the CPSC does not seem to have done this. Variance estimates are not currently part of the NEISS News, published by the U.S. Consumer Product Safety Commission.

The sampling plan used by the NEISS is described by the CPSC as a "two stage" sampling plan. The detailed description of the sampling plan and estimation equations are presented in Appendix A. In two-stage sampling, there are typically two components in formulas for the variance estimates. One term comes from each stage of the sampling.[1] In the sampling plan used by the NEISS, there is no within hospital variation, since a census of in-scope emergency room reports is collected. (There could exist some minor errors associated with censuses, such as omission or incomplete coverage, however.) There should still be two components to the variance, however, one term arising from the sampling of the hospitals within the selected strata and one term arising from the sampling of the strata from the super-strata. This does not appear to have been realized. In fact, the first stage--the sampling of the strata from the super-strata--seems to preclude the estimation of that component, since only one stratum was selected from each

[1]Cochran, W.G. (1963). Sampling Techniques, pp. 276 ff.

superstratum. It is, of course, impossible to form any estimate of variability from a sample of size one. This is an aspect of the sample design which we would hope to see altered as soon as a new sample is implemented. Our recommendation would be that that should be done as soon after the 1980 census as practicable.

As stated, it appears that one component of the sampling variance is not estimable. Thus, the estimate presented in Appendix A represents the variability among the second stage only. Further, the expressions presented there are unnecessarily complex. Apparently the author neglected to realize that while the variability of the numbers $\{q_{hij}Y_{hij}\}$ might be less than that of the numbers $\{Y_{hij}\}$, when calculating a variance of numbers multiplied by a constant, the variance is multiplied by the square of that constant. Thus any supposed gain in reduced variability would be eliminated. As is shown in Appendix B, the formula presented for the variance should be:

$$\hat{V}ar(\hat{Y}_{hi}) = \frac{1}{r_{hi}-1}\left[\sum_{j=1}^{r_{hi}} (W_{hij}\hat{Y}_{hij})^2 - \frac{\{\sum_{j=1}^{r_{hi}} W_{hij}Y_{hij}\}^2}{r_{hi}}\right] \quad (1)$$

and

$$\hat{V}ar\ (\hat{Y}) = \sum_{h=1}^{13} \sum_{j=1}^{3} \hat{V}ar(\hat{Y}_{hi}) \ . \quad (2)$$

Thus, the variance estimate is the sum of the within-strata variances. The within-stratum variance estimate is given above in (1). It is perhaps simplest to consider the within-stratum variance as arising from simply forming r_{hi} numbers for the $(hi)^{th}$ stratum, consisting of $\{W_{hij}Y_{hij}\}$, and then simply computing the sample variance among that sample of r_{hi} numbers.

As mentioned before, the estimate of error most often used in sampling work is the mean square error (or root mean square error). The mean square error is composed of two terms:

$$MSE = Variance + Bias^2$$

The estimation of the bias is quite difficult. Contributions to the bias term come from the aging of the sample, missing data, incomplete coverage, biases resulting from differential recording of associated products on the ER record, etc. It is uncertain to what extent the CPSC has considered these as sources of error. The aging of the hospital sample has been considered by CPSC and thought to introduce only a minor amount of error--2 to 3%--but it is not clear whether the percent refers to percent of the variance, percent of the standard deviation, or, most likely, to percent of the annual total.

Sample designs which are primarily concerned with the sampling variance may result in large bias errors if they are difficult to implement. Thus it is important to consider both types of error. Certainly if the bias became as large as the variance, more effort should be spent on the implementation to reduce the substitution rate, improve basic data control, etc., to reduce the bias. In this sense, a bias term approximately the same as the variance might be a reasonable compromise between sampling efficiency and practical implementation. If this is assumed, then the MSE would be about two times the variance, and the root mean square error would be increased (over the standard deviation) by a factor of 1.414.

As part of the standard summary analysis which we have developed and which is presented in Section V, we have used the formulas presented above to estimate the variance of the estimated total number of injuries per year. These values are underestimates of the actual errors, since they do not take the two-stage nature of the sample into account, and perhaps more importantly, they do not include any contribution for the bias term. Nevertheless, they are useful to give some idea of the magnitude of sampling errors arising in the estimation of the national totals for a product. For bicycles the standard deviations are estimated to be about 22,000 for each year's total. This compares to an estimated total of about 450,000. This gives a coefficient of variation of about 5%. If $\pm 2\sigma$ is used for a 95% confidence interval, then the interval for bicycle injuries is thus approximately from 400,000 to 500,000 injuries per year.

If standard deviations for the hazard index numbers were also calculated, they would tend to be even larger, since multiplying a value by a positive constant also multiplies the estimated standard deviation by that same constant. And the hazard index is essentially a weighted sum of constants times the estimated total injuries in each of several classes. It seems likely that the resulting confidence intervals would generally overlap for several products. Thus, it might well be that ten or fifteen products would have hazard intervals that overlapped, and in that sense should be grouped together. In point of fact such a grouping of products might better suit the aims of the CPSC than the current list does. The groupings might also be less suceptible to misreporting or misuses by the media. The actual variance and standard deviation estimates for bicycles are reported in Table 18, along with other aspects of the bicycle injury distribution.

E. Product Comparisons

The use of the NEISS to make product comparisons implicit in the hazard index is currently still on a somewhat shaky basis. In part this is due to the lack of exposure information which would allow the development of a hazard index based on risk rather than simply on the type and extent of use of a product. Current comparisons among products contain to a large extent the differences in the numbers of persons using the product and differences in the amount of use by individuals of the products. Consequently, the differences in the hazard index may be as much a function of different patterns of use and different persons who use a product as of different hazards.

To some extent products will always remain noncomparable. It will probably never be completely possible to meaningfully compare bath tubs with bicycles or bottle sterilizers with sliding glass doors. The uses of the products are simply too different. What must be kept in mind is the alternative to any given product. If a bicycle is used as a means of personal transportation, the alternative may be a motor scooter, motorcycle, or car. Aside from energy considerations, the alternatives may be more dangerous or illegal. If a bicycle is used simply as a toy, the alternative would be some other type of play--either with another toy or some other form of play activity such as climbing trees or playing in a construction area. In order to adequately determine whether a particular product poses an unreasonable risk, the risk of the alternative activity must also be considered.

Fortunately, it is not necessary or even desirable to compare products of varying uses. What is intended is to first identify products where large numbers of accidental injuries occur--whether through excessive risk or merely through widespread use--and then determine through in-depth investigations to what extent the accidental injuries could have been prevented by better product design, better maintenance, or more careful use. When this has been determined, appropriate remedial action may be taken to improve the safety of the particular product.

One serious problem in the process that has arisen is that the product hazard index has been widely publicized by the media. Even though a particular product has had a large number of in-depth investigations conducted by the CPSC and may no longer be of primary interest (as measured by the actual number of in-depth investigations sought by the CPSC), the media continue to report the product as being the most hazardous. This is a most regrettable misuse of the NEISS data, which cannot be prevented by the CPSC. The effect on the industries involved could probably be alleviated, however, if the CPSC were to report publicly that the products ranking high on the hazard index had been investigated in-depth and what corrective action--if any--had been found necessary and was being implemented. Such a statement, together with a statement that the products would be re-investigated at a later date to evaluate the effect of the remedial action and to ensure that the problem had been solved, could alleviate the adverse publicity which currently attends the products on the top of the hazard index. Such an attempt by the CPSC would be particularly useful since at present the hazard index is so closely tied to the degree of use of the product.

One of the current problems with the hazard index in its present form is its misuse by the media. This is particularly of concern to manufacturers since the ranking of the products depends to a considerable extent on the particular severity scores and "age adjustment" used. From the point of view of the CPSC and in terms of the proper use of the hazard index, shifts in rankings of several places are minor. Products are viewed as groups and selected for evaluation not only from their rank on the hazard index, but also from the results of previous in-depth investigations, consumer complaints, and other sources. However, from the point of view of the affected industries, changes in rank can be quite important. The media tend to focus on "public enemy number one" or the five worst, etc. Thus a shift of a few places in rank may result in an entirely different public relations image of a product.

An additional aspect of product comparisons relates to time. The hazard index has changed considerably in the past three years and will in all probability continue to evolve. It will certainly change markedly when exposure data are brought into consideration as a denominator. These changes in general tend to improve the utility of the index. However, they do cause a problem in that a particular product cannot be compared over time since the index is radically different from year to year. One of the major uses of the NEISS data should be to develop the potential to determine time trends in the injuries associated with a product. These, if properly corrected for seasonal effects and population changes, etc., can aid in determining whether a product is becoming safer, remaining about the same, or whether it has become more hazardous through some change in design or construction. Such time trend information is also of crucial importance in evaluating the effect of product safety standards developed and implemented by the CPSC. There is clearly a responsibility implied to determine whether a standard was sufficiently efficacious to justify any increased cost or inconvenience--or to determine whether additional steps for product safety are necessary.

Consequently, whenever the hazard index is changed, the new version should be used to re-calculate the index for the data from the previous years. It would also be useful for the CPSC to publish such year-to-year comparisons to provide the public with a meaningful comparison and to provide industry with information as to whether their attempts to improve the safety and usefulness of their product were being successful.

The CPSC Bureau of Epidemiology maintains that the data resulting from the NEISS collection system is used only as a guide in establishing their priorities for the conduct of in-depth investigations. The higher ranking products on the hazard index generated by NEISS data would be more likely candidates for further study into the specific causes of accidents and whether they are product related or not. The establishing of priorities for assignment of in-depth investigations will, of course, depend on other factors: how many cases for a product are already on file, how old the information is, and whether the cases have indicated that product defects have been identified as a significant factor in accident causation.

The ranking of products by the NEISS hazard index is often, however, presented to the public as an indication of the relative dangers inherent in various consumer products. It must be mentioned that this presentation is done not by CPSC but by independent journalists who have free access to the Commission's data under their open-door policy. A different index might be created which reflects the Commission's resource allocation decisions more directly than the single input to that process represented by the NEISS ranking of products. This index would rank products according to the number of cases assigned by CPSC for in-depth investigation, thereby ranking the Commission's concern over the relative dangers of the various products. Such an index should be used with the same caution as the NEISS hazard rankings remembering that it represents only a partial explanation of CPSC's behavior. It represents a combination of presently construed dangers as indicated by current information, with the status of investigative data already available in the Commission's files.

With regard to indices in general, a statement made about mathematical models would apply equally well: They are to be used, not to be believed. Unfortunately some people, for whatever reason, may like to attribute more meaning to an index than it deserves in light of the basic data from which it is constructed. It should be incumbent on the agencies which generate indices, therefore, to prevent insofar as possible, misinterpretations the general public might make of the numerical quantities involved. We applaud the recent decision of CPSC to rescale its hazard index. A number such as 863,200 might reasonably be assumed to be a report of the number of injuries resulting from a particular product on the basis of NEISS investigations. The value 86.32, while having the same utility for the Commission, is less likely to lead to such misinterpretations.

V. BICYCLE ASSOCIATED ACCIDENTS

A. Subset of Causality Factors

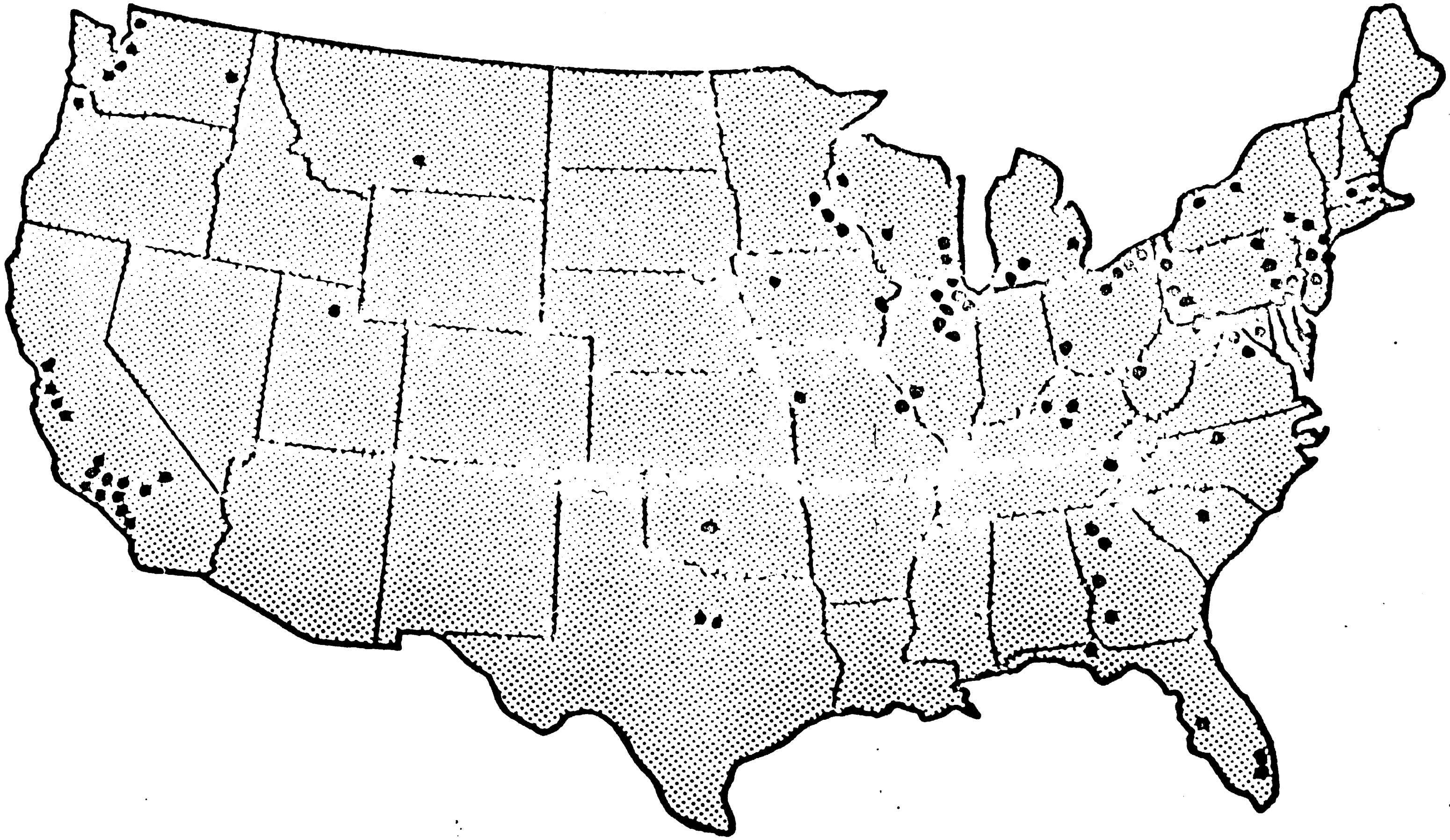
Several deficiencies in the sample design and the data collection and management of NEISS by the CPSC have been discussed in the preceding sections. A number of these deficiencies seem to apply with particular force to uses of the NEISS data in considering bicycle accidents.

As mentioned before, the error in the sampling scheme resulted in the selection of a complicated geographical subset of the U.S. as the population that was in fact sampled. This subset is roughly the drawing areas for the 2,300 hospitals which were not eliminated at the first stage. Since the elimination was based on probabilities which were proportional to population, the chosen subset is presumably of higher population density than that of the U.S. as a whole. In viewing a map with the NEISS hospitals located on it (Figure 6) it can be seen that large regions of the Midwest and West (other than the West Coast) have been excluded. Thus the regions which were in fact sampled would have been those with high traffic density, and could be expected to have rather more bicycle-traffic accidents than the more rural, small-town areas.

A great deal of the bicycle use in this country is by children. Accidents which result in minor injuries to

FIGURE 6

Location Of Hospitals In The NEISS Network



children are more likely to result in medical attention than are the same injuries to adults. As a consequence the ER rates will typically be more complete for younger children than for adults. Hence products which are used more by children than by adults may be expected to appear in somewhat excessive proportions. When coupled with the "age-adjustment" to severity suggested by CPSC, this appears to present a strong bias toward over-representation of young groups.

The calculation of the so-called "consumer product hazard index" and its publication by the CPSC does not appear to be an appropriate use of the surveillance data from NEISS. First, the NEISS data do not appear to be sufficiently reliable for such use to be made of them, due to the sampling and data collection inadequacies. Second, the index is based on an entirely arbitrary injury severity scale. Third, the index is based on mere association--there are not data to indicate the presence or degree of product fault in any of the injuries. Fourth, the index is simply a sum of raw data; the more popular and widely used a product is, the larger its score on this index will be. In order to construct a "hazard index" that indicates the relative risk of products, one must have and incorporate some type of exposure data.

An index of the kind used by the CPSC can be useful as an internal device to alert the CPSC to areas where there are substantial numbers of accidental injuries. However, to some extent it does this by pinpointing products which are widely used. It does not pinpoint products which are particularly hazardous to the user. In order to assess the hazard of a product, more information is needed. First, there is a need for exposure data. This need seems to have been somewhat slighted by CPSC in the past, but their future plans appear to include a serious effort in this direction. Secondly, there is a need for additional information on the types of injuries and their relation to the product in question. It is crucial to determine whether the accident and/or resulting injuries were the result of some cause unrelated to the product in question, or were in fact the result of the product. If they were actually related to the product, then the possibility of some action to reduce the risk or severity of such injuries is present. However, there is a need for sufficient information to determine whether the cause was due to a defect in design or construction, which would call for a construction standard; or to inadequate maintenance, which might call for an educational standard; or due to intentional misuse of the product, which might call for a warning standard, or perhaps for legislation dealing with the use of the product (enforce the prohibition against riding a bicycle the wrong way on a one-way street, for example).

In order to obtain data of sufficient detail to use in these aspects of the CPSC, in-depth investigations are necessary. The cases for such investigations must be selected and pursued in such a manner that the results can be related to the frequency estimates for the whole set of accidents. Thus, in order to recommend a bicycle standard, information should be available to indicate how it would help prevent injuries in certain types of accidents, and also what the frequency of such accidents is. It is impractical to write a standard which could apply on the average, to only one accident a year.

B. Motor Vehicle as Second Product

Table 16 shows the NEISS data for calendar 1974 broken down by injury category and by other product involvement. The separation of other products has been made by segregating motor vehicles and all others. We note that in only 2.7% of the reported cases is a motor vehicle found to be the second product involved, and conclude that motor vehicle/bicycle accidents do not comprise a significant proportion of all bicycle accidents. Furthermore the motor-vehicle-related accidents within the bicycle group do not appear to be any more serious than those involving either bicycles alone or other products.

In pursuing this conclusion we have looked at the computer files maintained by HSRI containing motor vehicle

TABLE 16. PROJECTED BICYCLE INJURIES FOR CALENDAR 1974 BY INJURY CATEGORY AND ASSOCIATED SECOND PRODUCT

	SECOND PRODUCT										
	None			Motor Vehicle			Other			Total	
Injury Category	Row %	Total Injury	Col %	Row %	Total Injury	Col %	Row %	Total Injury	Col %	Total Injury	% of Total
None	87.5	3,835	.9	2.2	94	.7	10.4	456	2.4	4,385	1.0
1	92.6	110,947	26.0	3.7	4,401	35.1	3.7	4,432	23.6	119,780	26.2
2	92.7	79,055	18.6	2.3	1,932	15.4	5.0	4,303	23.0	85,290	18.7
3	94.2	102,526	24.1	2.4	2,562	20.4	3.4	3,742	20.0	108,830	23.8
4	94.5	95,119	22.3	2.2	2,188	17.5	3.3	3,339	17.8	100.647	22.0
5	92.0	24,567	5.8	2.3	746	6.0	5.2	1,384	7.4	26,698	5.8
6	86.2	9,757	2.3	5.1	582	4.6	8.7	985	5.3	11,324	2.5
7	79.1	269	.1	10.7	37	.3	10.2	35	.2	340	.1
8	15.1	14	.0	0.0	-0-	0.0	84.9	76	.4	89	.0
Total		426,089			12,543			18,750		457,382	
% of Total			93.2			2.7			4.1		100.0

accident data. In none of the eight files examined did the frequency of bicycle accidents even approach that found in the NEISS data. The Texas 1974 5% sample file, for example, identifies 144 bicycle accidents out of a total of 21,058, or 0.68%.

The comparison between the automobile accident data banks and the NEISS reporting system is made more difficult by the incompatibility of the injury scales used. In none of the HSRI files was the injury reporting scheme sufficiently detailed to be comparable with that of the NEISS. Most of the reporting scales used are best approximated by the 5-point "police scale" (killed, A, B, C, and None). One scale approximates the 9-point "Abbreviated Injury Scale." Injuries in most cases, however, are grouped and it is difficult (if not impossible) to assign specific injuries or their severity to a particular participant in the accident, especially to a bicyclist.

VI. A STANDARD SUMMARY OF DATA

Although the NEISS data lack several data elements which would be useful, they nevertheless represent the most nearly nationally representative source of data extant in the U.S. which deal with in-scope product associated injuries treated at emergency rooms. As such it is appropriate that they be utilized to the fullest extent possible. In line with this, we have developed a standard summary or analysis of the data which can be obtained for a specific product code or for a selected grouping of product codes. This analysis is based on the national estimates available from weighting the observed injuries with the hospital weights. Naturally the summaries must perforce use the injury definitions set by the CPSC as well as the determinations of product association. This summary is presented for bicycles (code 1202), since that is the primary focus of this research. However, it may also be considered an example of the analysis which could be done for other products, for example, power mowers of specific types or of all types combined.

The single most interesting piece of information is probably the projected annual injuries associated with a given product, together with an estimate of the sampling error associated with that estimate. The estimated total number of bicycle-associated injury accidents treated in emergency rooms in the U.S. for the three calendar years 1972, 1973, and 1974 are presented below, with the estimated standard deviation of the total, estimated according to the formula presented in Appendix R. (It is to be recalled that this estimate is probably on the low side and does not include any estimate of the bias error.)

TABLE 22. ESTIMATED NATIONAL TOTAL BICYCLE-ASSOCIATED INJURIES

	1972	1973	1974
Estimated Total	334,100	419,920	457,380
Estimated Standard Deviation	17,649	20,810	23,296

It will be noticed that there is an increase in the total for each year. This is thought to be due to a combination of factors: population increase and population shift to younger ages, increased use of bicycles as a means of transportation, more general acceptance of the bicycle as a vehicle for recreation among older age groups, more interest and increased reporting of bicycle association, etc. Although no firm conclusions can be drawn from three years, the increase from 1972 to 1973 seems unusually large and may be due to a reporting artifact in the NEISS caused by the first year of data collection resulting in an unusually low estimate for 1972. Alternatively, it could reflect a boom in bicycle sales and use in 1973. Probably some combination of these two factors plus some other is responsible.

The injuries reported to the NEISS are classified by body part and by diagnosis according to the codes presented in Table 17a. Table 17 presents the estimated number of each injury diagnosis and boyd part combination for bicycle-associated injuries for calendar 1974. Inspection of Table 17 reveals that four diagnoses accounted for about 93% of the injuries. These were: contusion or abrasion (34%), fracture (14%), laceration (36%), and strain or sprain (9%). Similarly, four general body regions accounted for about 86% of the injuries. These were: head (13%), face (21%), lower arms, hands, and fingers (24%), and lower legs, feet, and toes (28%).

Some indication of the population injured may be had by considering the age and sex distribution of the injured persons. The estimated number of individuals injured in each age and sex category is presented in Table 19. Approximately twice as many males were involved as females, while about 77% of the injuries occurred to persons under the age of 15 years. The age and sex distribution of bicycle-associated injuries in 1974 is shown in Figure 7.

The CPSC has developed a severity scale for injuries which consists of collecting each of the possible body part and diagnosis categories into one of nine categories according to severity. These severity classifications are used by the CPSC in determining a frequency-severity-index upon

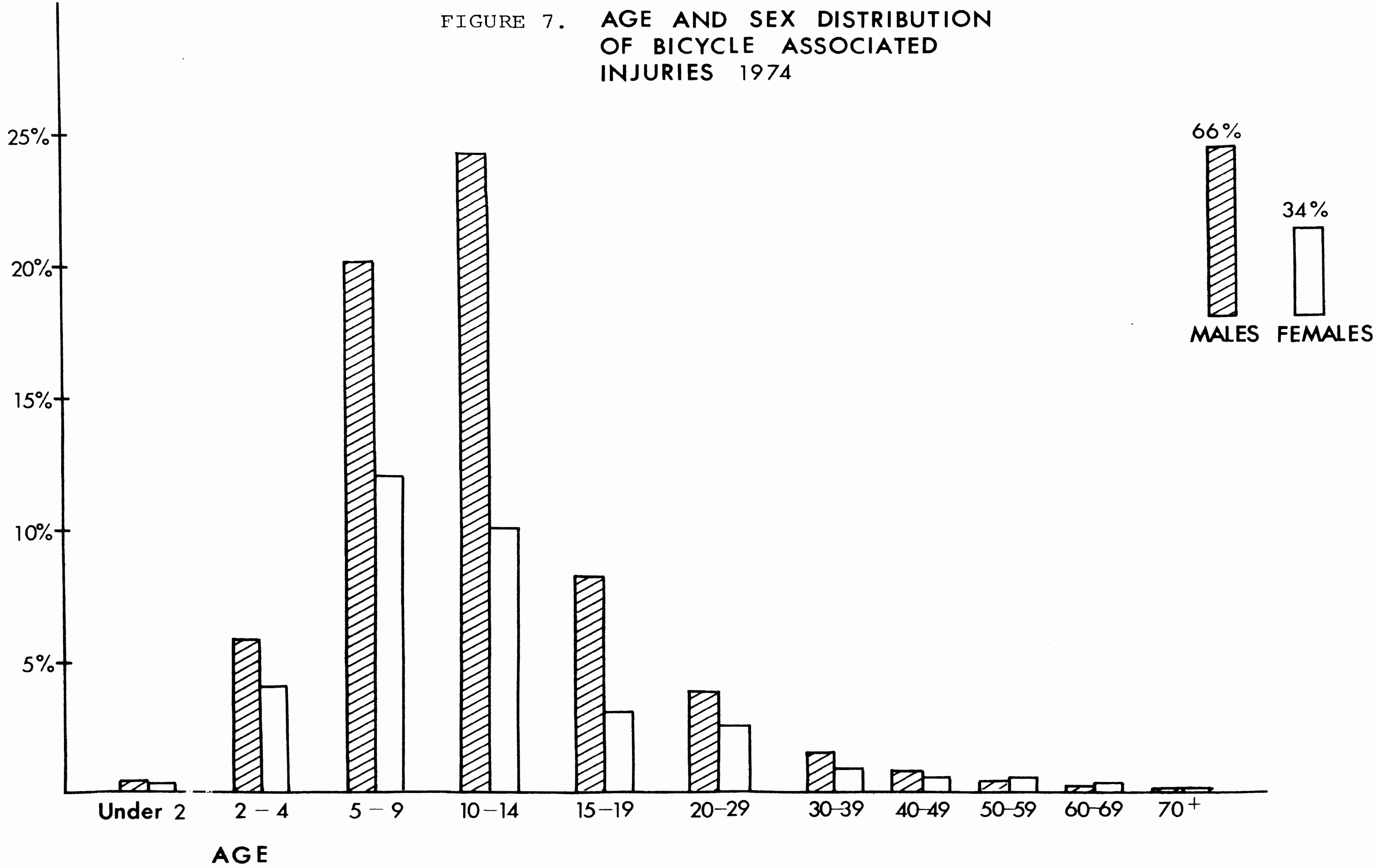
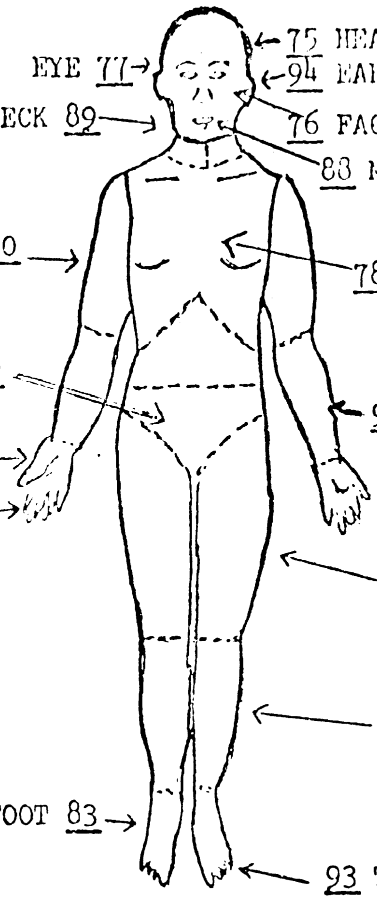


FIGURE 7. AGE AND SEX DISTRIBUTION OF BICYCLE ASSOCIATED INJURIES 1974

TABLE 17a. BODY PART AND INJURY DIAGNOSIS



INJURY DIAGNOSIS	CODE
Amputation	50
Anoxia	65
Avulsion	72
Burns (not specified)	47
Burns (scald from hot liquids)	48
Burns (thermal)	51
Burns (chemical, caustics, etc.)	49
Cell damage by radiation, except thermal (radiation burns by ultraviolet, x-rays, radioactive materials, etc.)	73
Concussion	52
Contusions/Abrasions	53
Crushing	54
Dermatitis, Conjuntivitis	74
Dislocation	55
Electric Shock	67
Foreign Body	56
Fracture	57
Hematoma	58
Internal Organ Injury	62
Laceration	59
Nerve Damage	61
Poisoning	68
Puncture	63
Strain or Sprain	64
Submersion (including drowning)	69
Other	71
Not stated	70
Ingested foreign object	4100
Aspirated foreign object	4200

BODY PART	CODE
Head	75
Ear	94
Eyeball	77
Face (including nose)	76
Mouth (lips, tongue, teeth)	88
Neck	89
Upper trunk (including shoulders)	78
Lower trunk (including hips)	79
Upper arm	80
Lower arm (including wrist and elbow)	90
Hand	82
Finger	92
Upper leg	81
Lower leg (including knee and ankle)	91
Foot	83
Toe	93
25-50% of Body	84
All parts of body	85
Other	86
Not stated	87

TABLE 17. 1974 ESTIMATED NATIONAL DISTRIBUTION OF BICYCLE INJURIES BY BODY PART AND DIAGNOSIS.

Body Part

Diagnosis	Internal Ingested or Aspirated	Head	Ear	Eyeball	Face Includ-ing Nose	Mouth Including Lips, Tongue, Teeth	Neck	Upper Trunk Including Shoulders	Lower Trunk Including Hips	Upper Arm	Lower Arm Including Wrist & Elbow	Hand	Finger	Upper Leg	Lower Leg Including Knee & Ankle	Foot	Toe	25-50% of Body	All Parts of Body	Other	Not Stated	Row Total	Row %
Amputation					5								254				21					280	.06
Anoxia																							0
Avulsion		19			58	285		28		6	59	102	1377	51	299	348	741					3373	.7
Burns (Not Specified)											40					10	6				94	150	.03
Burns (Hot Liquid)															2							2	*
Burns (Thermal)												5		6	4	13						28	*
Burns (Chemical)											30											30	*
Cell Damage (Radiation)										9												9	*
Concussion		9792																				9792	2.1
Contusion, Abrasions		19159	144	337	22840	1953	304	14615	6803	1363	22820	5182	4026	3308	32579	14536	1889	2512	326		150	154846	33.6
Crushing					238				20		19	9	658	7	57	18	46					1072	.2
Dermatitis, Conjunctivitis				69	121	19					16		27		27							279	.06
dislocation					39	825		1284		256	1079	43	603	5	140	17						4291	.9
Electric Shock																						0	0
Foreign Body		11		137	64			19		15	36	172	61		26	27						568	.12
Fracture		1185			2570	584	235	9987	475	1643	24356	2436	5455	746	9365	2115	1576			10		62738	13.6
Hematoma		2330	22	156	1696	54		192	347	37	338	265	687	39	909	258	249	6				7585	1.6
Internal Organ		24	6	130			53		480											94	30	817	.1
Laceration		23314	1027	321	53940	11134	193	1665	3066	653	7090	5093	8347	4226	29131	10289	4530	641	82	6		164748	35.8
Nerve Damage			21																			21	*
Poisoning																						0	0
Puncture		565			288	31		53	72		148	210	82	160	395	359	41	27				2431	.5
Sprain/Strain							621	3556	1461	189	11225	1483	3612	275	12798	4618	2720	24	9			42591	9.2
Submersion																						0	0
Other									26													26	*
Not Stated		1169	94	94	11	94		322	107		659		230	94	584	565					130	4153	.9
Ingested Foreign Object																							0
Aspirated Foreign Object																							0
Column Total		57568	1314	1244	81870	14979	1406	31721	12857	4171	67915	15000	25419	8917	86316	33174	11819	3210	417	110	404	459830†	
Column %		13	.2	.2	17.8	3.2	.3	6.8	2.7	.9	14.7	3.2	5.5	1.9	18.7	7.2	2.5	.6	.09	.02	.08		

†Discrepencies in total projected figures from chart to chart are due to missing values for variables used in a particular stratification.
*Comprises less than .1% of total.

TABLE 18. ESTIMATED TOTAL BICYCLE INJURIES BY MONTH

Month	1972	1973	1974
January	3,580	7,158	15,700
February	3,940	7,371	14,353
March	10,510	18,909	24,356
April	28,899	32,611	43,180
May	40,443	45,540	55,010
June	47,100	63,887	64,245
July	56,016	72,666	73,323
August	63,496	69,797	69,210
September	46,697	51,897	48,732
October	21,194	31,058	26,053
November	8,439	11,233	15,188
December	6,781	7,794	8,034
Total	337,095	419,921	457,384
Estimated Standard Deviation of Total	17,650	20,810	23,300

TABLE 19. ESTIMATED NUMBER OF BICYCLE INJURIES BY AGE AND SEX FOR THE UNITED STATES, 1974

Age	Males	Females	Total	Row %
Under 2	2,019	1,193	3,212	.70
2-4	27,016	17,077	44,093	9.67
5-9	92,061	54,720	147,781	32.42
10-14	110,871	45,658	156,529	34.34
15-19	37,576	13,800	51,376	11.27
20-29	17,042	11,187	28,229	6.19
30-39	6,505	4,218	10,723	2.35
40-49	3,662	2,658	6,320	1.39
50-59	1,665	2,719	4,384	.96
60-64	493	558	1,051	.23
65-69	410	837	1,247	.27
70+	500	391	891	.20
Total	300,820	155,016	455,836	
Column %	65.99	34.01		100%

TABLE 20. ESTIMATED NUMBER OF BICYCLE INJURIES BY NEISS SEVERITY CATEGORY FOR 1972-1974

Severity	1972	%	1973	%	1974	%
0	5,986	1.8	2,712	.6	3,818	.8
1	78,232	23.1	105,080	25.1	119,651	26.2
2	79,161	23.4	92,658	22.1	85,073	18.7
3	68,372	20.2	86,288	20.6	108,585	23.8
4	73,135	21.6	95,020	22.7	100,372	22.0
5	18,875	5.6	22,356	5.3	26,591	5.8
6	9,305	2.7	13,150	3.1	11,316	2.5
7	686	.2	1,775	.4	339	.1
8	122	.04	245	.1	89	.02
Total	338,884		419,284		455,834	

SUMMARY OF SEVERITY INDEX

Category 7 = Category 6's who are hospitalized and deaths = Severity Value of 2516

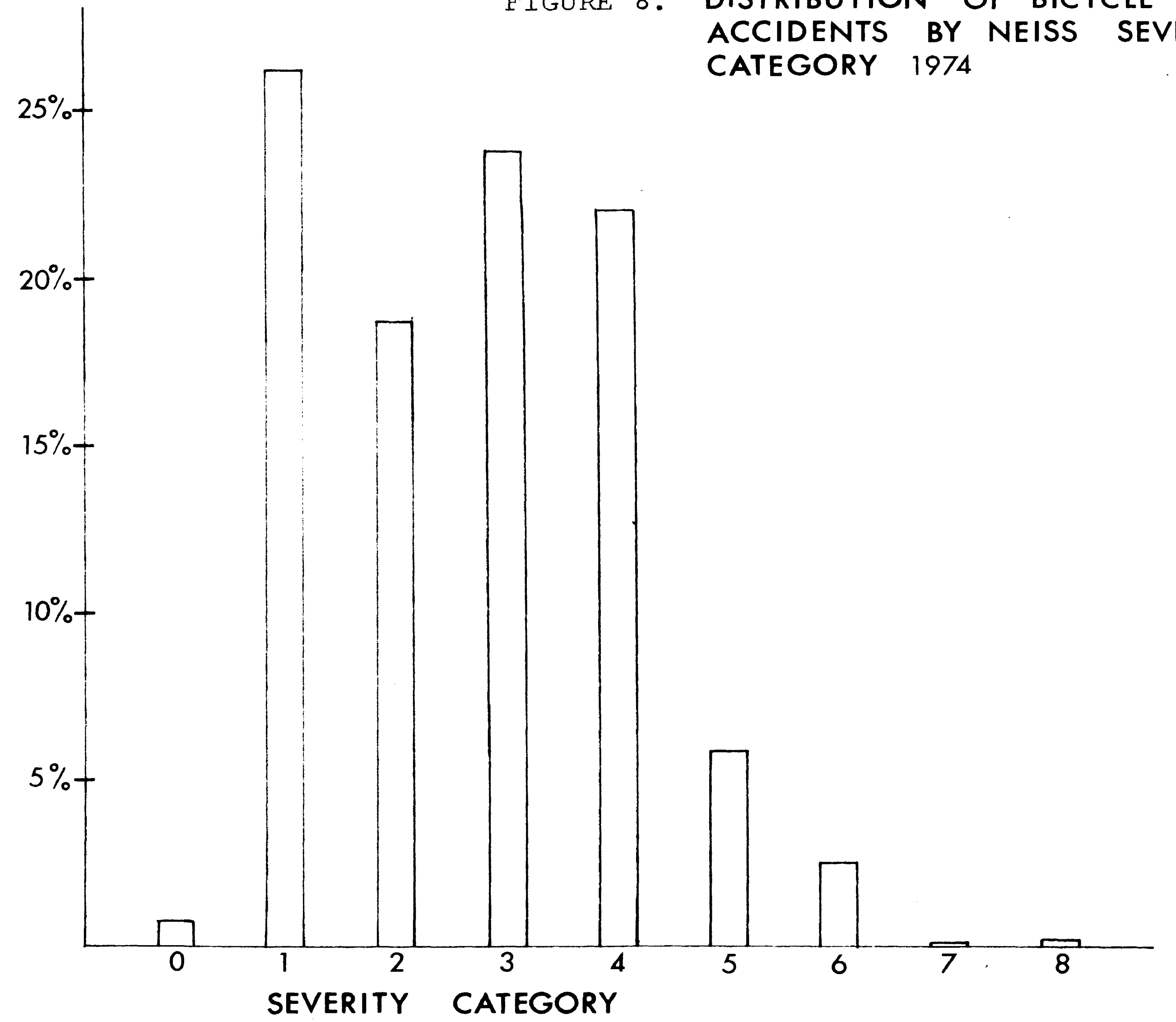
Diagnosis	Severity Category 6 Severity Value - 340	Severity Category 5 Severity Value - 81	Severity Category 4 Severity Value - 31	Severity Category 3 Severity Value - 17	Severity Category 2 Severity Value - 12	Severity Category 1 Severity Value - 10
Amputation	Any part of body					
Avulsion	25% of body +	head, eye, upper trunk	lower trunk	leg, arm, hand, foot, finger, toe	mouth, ear	
Burns	25% of body + or eye	all single body parts except finger, toe, ear			ear, finger, toe	
Cell Damage	25% of body +	head, face, eye, upper or lower trunk		leg, arm, hand, foot, finger, toe		
Concussion	25% of body +	head				
Contusion or Abrasion	25% of body +			head, upper trunk	ear, mouth, neck, eye, lower trunk	arm, leg, hand, foot, finger, toe
Crushing	head, arm, leg, trunk, foot, hand		finger, toe			
Dislocation	25% of body +	head, upper trunk	lower trunk, eye		arm, leg, hand, foot, finger, toe	
Foreign Body	25% of body +	head, upper trunk	lower trunk	mouth	arm, leg, hand, foot, finger, toe, eye	
Fracture	25% of body +	head, neck, upper and lower trunk	eye	arm, leg, hand, foot, finger, toe, mouth		
Hematoma	25% of body +	head, upper trunk	eye, lower trunk	arm, leg, hand, foot	finger, toe, ear, mouth, neck	
Internal Organ Injury	25% of body +	head, neck, upper or lower trunk	mouth, eye			
Laceration	25% of body +		head, eye, upper or lower trunk		arm, leg, hand, foot, finger, toe, ear	
Nerve Damage	25% of body +	all other body parts				
Puncture	25% of body +	head, face, upper trunk	eye or lower trunk		arm, leg, hand, foot, finger, toe, mouth	
Strain or Sprain	25% of body +			neck, upper trunk	lower trunk, eye	arm, leg, hand, foot, finger, toe, ear
	anoxia, electric shock, submersion	ingested or aspirated foreign object				
Dermatitis			25% of body +		head, face, eye, upper and lower trunk	arm, leg, hand, foot, finger, toe, ear

TABLE 21. NEISS INJURY MATRIX

which the product hazard index is based. These severity classes are presented in Table 21. Each of the severity classes is assigned a weight according to the scheme described in Section V. The product of this weight with the estimated number of injuries is then summed over all categories to obtain the frequency-severity index. (The hazard index also incorporates a weighting factor of 2.5 for accidents to persons under 15.) By referring to Table 20, one can determine the injury severity distribution of bicycle-associated accidents for the three years. This distribution can be combined with any desired weighting scheme to provide a summary of a linear combination of frequency and severity. Thus one is not restricted to the scheme used by the CPSC and described in Section V, but one can calculate a similar figure based on a different selection of the weights. Further, more meaningful summaries may be made in terms of the proportion of accidents of a given severity or less. As an example, one finds that nearly half (45.7% in 1974) of the injuries were of severity category two or less. It should be noted that severity category zero consists of persons with no injury but who were seen in an emergency room--presumably for examination to see whether they had been injured. This group is quite small. It ranges from 1.8% in 1972 to .6% in 1973 and .8% in 1974. For 1974, the distribution of bicycle-associated injuries by severity class is shown in Figure 8.

Many products are used more extensively during some months of the year than during others. This is particularly

FIGURE 8. DISTRIBUTION OF BICYCLE ASSOCIATED ACCIDENTS BY NEISS SEVERITY CATEGORY 1974



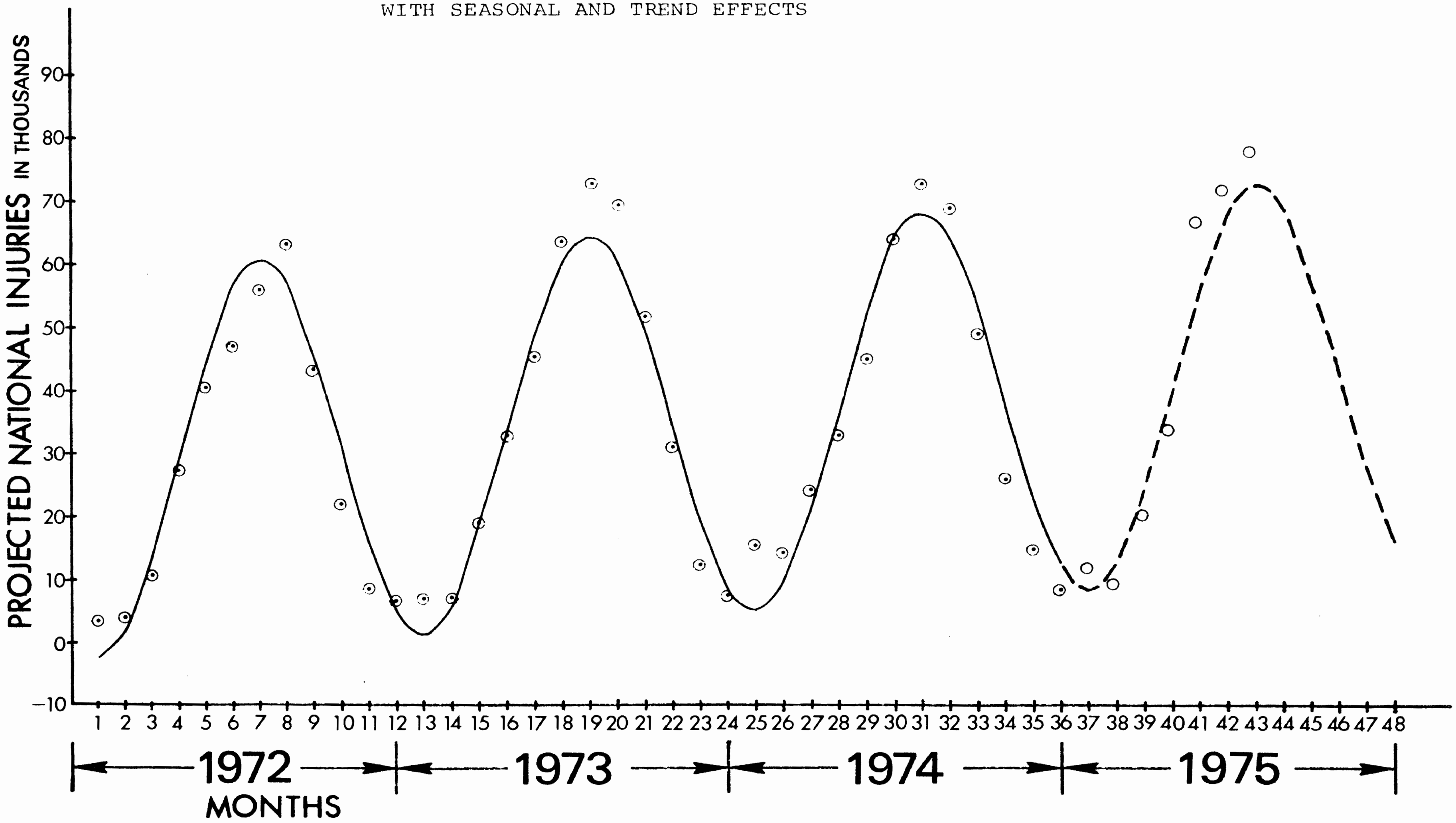
true of bicycles as well as of many other types of recreational equipment. Table 18 presents the estimated total bicycle-associated injuries by month for the three year period 1972 to 1974. This allows one to consider a time series analysis to estimate the seasonal component as well as a linear trend in the estimated total number of injuries. Two simple time series models have been fitted to the data of Table 18. The first model uses a sine function to represent the seasonal variation and a linear term to represent the secular trend. This secular trend is the gradual increase due to increasing population, gradual increase in traffic, and general increase in exposure. Using this model results in an explanation of 93.7% of the variation in the monthly injury projections over the three years. A plot of the injuries predicted by the model together with the projected injuries for the three year period is presented in Figure 4. The model has also been extrapolated to predict for the calendar year 1975.

The model used is given by the equation

$$Y = 27{,}760 + 318X + 30{,}750\,\text{sine}\left[\frac{2\pi(X-4)}{12}\right],$$

where X is the number of the month, starting with January, 1972 as 1, continuing on to December of 1974 as 36, and where Y is the predicted national total of injuries for the given month. Several points may be noticed about

FIGURE 4. ESTIMATED MONTHLY INJURIES MODEL WITH SEASONAL AND TREND EFFECTS

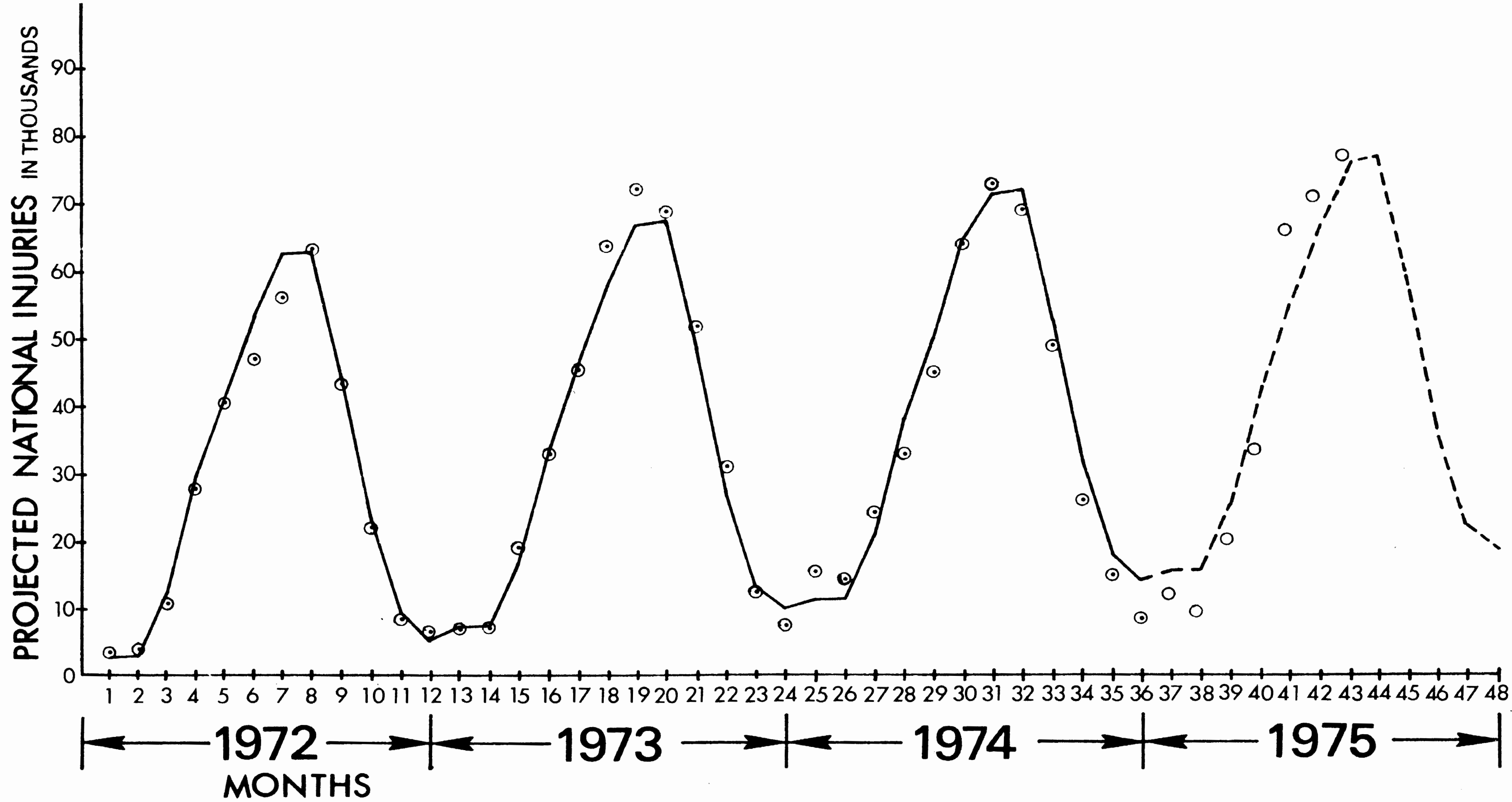


the model. The first is that the sine curve seems to be slightly out of phase for 1972. To some extent this might be corrected by using the data to estimate the phase of the sine function rather than using the nearest integer. However, it appears that the data for 1972 differ from those of the other two years in that a substantially smaller total of injuries were estimated and the peak came in a different month. Injuries seemed to be unusually low for the first six months of 1972. This is probably a peculariarity of 1972 and not a regular feature of the data. It could, in particular, be due to under-reporting in the startup of collection of data by the NEISS. Secondly, the model tends to underestimate the peak in the summer months, while also underestimating the injuries in January. The projected injuries show a small "bump" in January, which is probably due to the large number of bicycles which are received as Christmas presents and which are then tried out at the first opportunity even though the weather is not ideal for first riding a bicycle or first riding a new and larger bicycle. There may also be a similar bump after school gets out in the summer--usually the middle to latter part of June. This may account for the failure of the model to estimate quite enough injuries at these two times. A third point is that the model estimated a negative number of injuries for January of 1972. This points out that the model should not be used to extrapolate backwards in time. It may also be due somewhat to the relative low number in injuries in 1972,resulting in a somewhat too large value for the coefficient of the sine term. The one negative estimate is not regarded as a particularly serious defect in the model but it does point out that the model cannot be used uncritically.

A second model was fitted using monthly means to estimate the seasonal component and a linear term to represent the secular trend. A plot of the values estimated by this model together with the data is presented in Figure 5. The model is also used to extrapolate to 1975. The percent of the variation in the monthly injury projections explained by this model is 97.9%, so it is slightly better than the previous one in this regard. On the other hand, the previous model estimated only three parameters from the 36 data points, whereas this one estimates 14 parameters from the 36 data points. Thus the model is approaching saturation and little is gained in the second model over the first in terms of accuracy of prediction. The second model does incorporate the Christmas gift phenomonen and it does predict the summer peak somewhat better than the simpler model. This second model may be represented by the equation

$$Y = -7045 + 381X + M_i ,$$

FIGURE 5. ESTIMATED MONTHLY INJURIES--MODEL WITH SEASONAL, TREND, AND REGULAR EFFECTS



where Y is the predicted number of injuries for the month, X is the number of the month, beginning with 1 for January, 1972, and M_i represents the monthly adjustment as given below.

January	8,813
February	8,555
March	17,925
April	34,897
May	46,998
June	58,411
July	67,335
August	67,501
September	48,108
October	26,102
November	11,620
December	7,537

These are not the only models which could be fitted. They provide examples of models which predict the injuries reasonably well. Some kind of models like these would need to be used to estimate the effect of any new standards which were introduced, or to conclude that a change of some sort had significantly changed the injury picture. This is to say that an increase in injuries associated with any product should not be viewed as an increase in the risk associated with that product unless it represents a marked departure from the general trend as represented by some type of a time series model.

VII. LITERATURE SURVEY

Several studies in the last five to six years have attempted to draw conclusions about the nature of bicycle accidents occurring, the characteristics of the population to which they occur, and/or the properties of bicycles involved in accidents. They have ranged in scope from a single city's children to a state's bicycle-motor vehicle accidents over three years to a survey involving children in 120 nationally distributed school districts. In the last year there has been an apparent surge of interest in the bicycle with a corresponding rise in report production. Even the E.P.A. has issued an overview.

F. J. Vilardo and J. H. Anderson (1969) gathered exposure and accident data on 20,000-plus children in grades 2 through 8 in Arlington Heights, Illinois; Ann Arbor, Michigan; East Baton Rouge Parish, Louisiana; Los Angeles, California; and the State of Delaware. The sex and age distributions of all bicyclists and the approximately 4,000 experiencers of bicycle accidents were fairly similar. The percent distribution of ostensible accident causes is shown in Table (a)-2. Vilardo and Anderson observed that boys had more motor-vehicle-involved accidents and girls more falls (riders of girl's bicycles experienced "significantly more knee injuries"). In approximately 70 percent of the accidents the bicycle had been used as a toy, in 23 percent as a general transportation vehicle, and in seven percent as a transportation vehicle specifically to or from school. While 38 percent of exposure respondents claimed night-time riding, three percent of accidents occurred then.

E. Brezina and M. Kramer (October, 1970) looked at 275 bicyclists (5-14 years, 93 percent males) in reportable bicycle/motor-vehicle collisions in the ten-month school year 1969-70 and a randomly chosen comparison group of 1,000-plus eight- to 13-year-old male bicycle owners. According to the investigators, risk factors "related primarily to the bicyclist's comprehension of risk inherent in different roadway environments, prediction of traffic movement, and control of the bicycle...Among the bicycle factors examined, the adjustment of the bicycle to the rider has the most significant influence on the risk of collision, especially among young riders." Relative risk of collision involvement increased as bicyclist ground-level clearance increased. Approximately half the collisions involved standard configuration bicycles, while 35 percent of the comparison group owned such bicycles. The investigators, lacking use information by configuration type, were reluctant to draw firm conclusions regarding relative risk but they suggested "among 11- to 13-year-olds a lower collision involvement rate...is found for high-rise bicycles" compared with standard configuration bicycles. Further, they suggested younger riders of hand-brake-equipped bicycles are more at risk than those of bicycles with foot brakes (note that Rice and Roland (April, 1970), studying the dynamics of bicycle design, suggest front (hand) brakes have "hazard potential"). Fifty-five percent of 259 collisions involved bicycles crossing some flow of traffic; 57 percent occurred within one block of home.

In February, 1969, P. A. Waller and D. W. Reinfurt published an investigation of some 2,400 fatal and non-fatal bicycle/motor-vehicle accidents reported in North Carolina, July, 1965, to June, 1968, relative to 1966 motor vehicle accidents. They analyzed the physical circumstances surrounding the accident, driver and bicyclist characteristics, and vehicle (car) and road characteristics. The authors summarize: "It appears that the typical bicycle accident occurs in clear dry weather during daylight hours...The cyclist is usually a young male between 10 and 14 years of age who apparently emerges unexpectedly from...(an) intersection of some sort. Fatal bicycle accidents appear to be associated with the older bicycle rider." About 25 percent of bicycle fatalities occurred at night on unlighted roads, but only six percent of all bicycle accidents occurred under such conditions.

A four-month study of ownership, use, and injury patterns among three- to 12-year-olds in the Burlington, Vermont area (J. A. Waller, 1970) was intended to determine if standard and high-rise bicycles differ in associated rate and/or severity of injuries. With 104 injuries (97 percent of all occurring) and 6,200 comparison cases, no differences were observed between the two bicycle styles. A percent distribution of accident causes is shown in Table 23-4. While about 30 percent of owners stated they rode after dark, no accidents occurred then. Eight accidents involved contact with an automobile; 69 percent occurred within one block of home. In the range of five to 12 years, of 6,185 owners a four-month injury rate of 11.2 per 1,000 owners (based on a total of 69 injuries) was calculated.

During approximately the same period as the Vermont study, exposure and accident data (the former from a school sample, the latter from the sample plus emergency room and police department records) for children five to 18 were gathered in Raleigh, North Carolina. E. A. Pascarella (1971) found no real differences in frequency or category of accident associated with bicycle style (a breakdown of the latter may be found in Table 23-1. Accident subjects tended to be younger and to have fewer years riding experience than those without accidents. Bicycle age and condition and passenger-carrying status did not appear to

TABLE 23. BICYCLE ACCIDENTS BY TYPE.

	1. Pascarella (1971)			2. Vilardo (1969)			3. Chlapecka (1975)			
	All Types	High rise	Std.	All Types	Middle weight	High-rise	All Types	Middle-weight	High-rise	Light-weight
N	218			3,952			2,884			
All Types	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%	100.0%
Bicycle Struck Car	2.8	1.9	3.5	7.9	8.4	6.0	5.5	5.4	4.8	10.5
Car Struck Bicycle	6.4	3.7	8.2	5.3	5.3	4.4				
Bicycle Struck Fixed Object	4.1	3.7	5.9	15.1	14.5	16.0	3.3	3.2	2.9	6.3
Bicycle Struck Bicycle	7.3	8.3	4.7	18.3	17.4	19.3	30.1	28.8	30.5	30.5
Fall	72.1	75.0	69.4	53.3	54.3	54.3	11.9	11.7	12.0	11.6
Other	7.3	7.4	8.3	---	---	---	49.2	50.0	49.8	41.1

	4. J. A. Waller			5. NEISS	
	All Types	High-rise	Std.	All Types	
N	104			119	
All Types	100.0%	100.0%	100.0%	100.0%	
Car Involved	9.6	2.0	14.0	---	
Hit Obstruction	17.0	15.0	19.0	---	
Loss of Control	20.0	22.0	18.0	63.2	
Foot Caught	11.0	10.0	11.0	10.8	Body Entanglement
Bike Broke	2.0	2.0	3.0		
Rode too Close to Stationery Object	8.7	10.0	8.0	16.7	Mechanical & Structural
Other	31.7	39.0	27.0	9.2	

influence risk...An overall accident rate of 1.58 per 1,000 miles (based on 60,109 miles, 95 accidents and 397 owners) was calculated. It was further estimated that 15.7 percent of accidents would result in no injury, 76.3 percent in mild injury (first aid only), and eight percent in physician-, emergency room-, or other hospital-treated injury. (Contrast this distribution to the Market Facts, Inc. survey (1970) cited by NEISS that says 82 percent of injuries are treated professionally.) Pascarella uses ownership, mileage, and accident rates to project approximately 9.46 million accidents occurring annually in the United States, with about 760,000 requiring medical attention. Note that this estimate, made in 1971, is nearly twice that made by NEISS (453,000) in 1973.

NEISS conclusions about the association of accidents with any particular aspect of bicycle style, structure, or condition are made primarily from their "in-depth" investigations rather than reports on all accident cases arriving at emergency rooms. These in-depth cases have included referred death certificate information as well as any very severe (category 6, 7, or 8) injuries and cases that look interesting or are called to someone's attention. The surveillance data give frequencies of injury types and of age and sex groups. For example, of 11,403 injuries during the period July, 1972 - March, 1973, 13 percent were fractures (J. A. Waller reported 20 percent fractures out of 104 injuries in Vermont). Using NEISS in-depth cases, a percent distribution of accident causes is shown in Table 23-5. The age and sex distributions of non-fatal (in-depth) cases and of total input records from the May, 1973, Staff Analysis are virtually the same.

During 1974, A. Williams of the Insurance Institute for Highway Safety examined 888 Maryland-reported injury-producing bicycle/motor-vehicle collisions for operator characteristics that might suggest responsibility. Ninety-nine percent of involved bicyclists, 84 percent (27) of their passengers, and one percent of involved motorists were injured. Approximately 75 percent of involved bicyclists were under 15 years; seven percent of collisions occurred after dark. Probable responsibility for collision appeared strongly related to bicyclist age and not to motorist age. Light conditions, location, and maneuvers were related to bicyclist age as well as to probable responsibility. Probable responsibility had been designated as the bicyclist's in 78 percent of the collisions investigated.

C. L. Lefler summarized the results of a 1973 Santa Barbara, California, study as presented in two analyses (Popish and Lytel, June, 1973 and Cross and deMille, June, 1973) in a paper given at a bicycle/pedestrian facilities design conference in 1974 (MAUDEP, p.279). Factors identified as accident-related but not merely bicyclist-related in these analyses included visual obstruction and bicyclist visibility, bicyclist age, bicyclist's route, traffic violations, hazard recognition, inattention, misjudgment, operator and victim attitudes. Interesting in this paper is the argument for a set of rules based on children's cognitive development, including a minimum age for bicycle use.

Coincident with their introduction of a bicycle law enforcement program in November, 1972, the California Highway Patrol (CHP) carried out a study of California bicycle accidents relative to those people involved and to the population of all bicycles ridden. In examining 1969-1973 accidents by applying regression techniques (no equations are provided), the CHP concluded there was a significant drop in accident occurrences from 1972 to 1973. Although the relationship was statistically significant in only one instance, several relationships appeared to hold between increased citations and decreases in collisions produced by the corresponding violation. Repeat violators and all violators showed approximately the same distributions of violations. Using a sample of bicycles found on school grounds and campuses as a comparison group, the CHP decided that, on the whole, accident-involved and all bicycles were similarly equipped, with a slightly greater portion of comparison bicycles equipped with chain guards. Approximately 61 percent of all fatal or injured bicyclist victims were 14 years old or under and 45 percent of motor vehicle drivers involved in bicycle accidents were under 30 (vs. 51 percent of those involved in any accident). Of a 214-accident-bicycle sample, about 16 percent had been used for riding to or from school.

A recent (1975) National Safety Council survey was made of 23,699 elementary-school-aged bicyclists from 120 schools across the United States. It attempts to some extent to generalize the studies of Vilardo, Pascarella, and others to the nation by starting with a national sample and by incorporating NEISS data into its reported accident group. The structure of the study is similar to that of Pascarella with data gathered on accidents in the previous year and in two to five years previous (to validate the 1969 Vilardo and Anderson study). One interesting point was that, of the approximately 16 percent of injuries that were treated at a hospital or by a doctor, 38 percent were scrapes, cuts, bruises, or sprains (vs. 87 percent of all injuries). The authors projected that, nationally, "336,000 bicyclists aged five to 14 sought treatment at a hospital including those who were admitted as in-patients" from June, 1971 through May, 1972. A percent distribution of accident cause by bicycle style is shown in Table 23-3. Based on the study group (N = 9,425), an overall accident rate for the total group of 0.72 per 1,000 miles driven was calculated. Accident rates per 1,000 miles driven were highest for 12- to 13-year-old males with fewer than two years experience.

TABLE 24. BREAKDOWN OF "MECHANICAL AND STRUCTURAL"

	Pascarella (1971)	NEISS (1973)	Chlapecka (1975)
N	30	20	758 (out of 3005)
All types	100.0%	100.0%	100.0%
Brake Failure	20.0	35.0	26.1
Wheel and/or Handlebars Unstable or Disengaged	50.0	35.0	30.9
Other	30.0	30.0	43.0

VIII. CONCLUSIONS AND RECOMMENDATIONS

The NEISS is the best source of accident injury data (for injuries treated in emergency rooms which are associated with in-scope products) for the United States today. There remain some features of the system which could be improved, and we hope that the recommendations made here will be useful in the continuing evolution of the system.

The data which are available from the NEISS should continue to improve in quality and usefulness, and it is our hope that the conclusions and recommendations below will be useful in improving the inferences from the data at present and in suggesting analyses which may make the data more useful and understandable to more people. It is also our hope that a number of doubtful or inappropriate uses or interpretations of the data which have been published can be improved.

A. Sample Design

The sample design seems unnecessarily complex and difficult to comprehend. It is therefore not likely to gain acceptance or belief in the manufacturing community or among professionals. No attempt had been made to determine the possible effects of substitutions in the primary sampling units. This leaves the question of bias error quite open.

The NEISS is a sampling of existing records. Thus there is no control over primary data collection. As a result, inconsistencies among institutions are to be expected. More seriously, there are likely to be large differences in reporting product association among different products. The lack of primary data collection control is the most serious deficiency in the NEISS. Population shifts and growth have taken place since the data used to determine the sample (the 1960 census) were collected. This aging of the sample will always be somewhat of a problem. The error that this introduces is not large currently, but will continue to grow.

In view of the conclusions above and the practical considerations of drawing a large sample, we recommend that a new sample design be developed, and that it be based on the data from the 1980 census and implemented as soon after those data become available as practical. We also recommend that the persons collecting the primary data on each emergency room visit should be supported by the CPSC and trained by them so that control of the primary data collection can be established. The assumption of this primary data collection responsibility by the CPSC could be implemented in the present system and would result in a marked improvement in the NEISS. We recommend that it be implemented as soon as possible.

B. Reporting of the Data

Any national estimates, whether based on a census or on a national sample, should be accompanied by estimates of the error--sampling error and bias--which may occur. To date the CPSC has not provided any error estimates. There seems to be some question as to whether the sample design actually implemented admits to exact variance estimation. The estimation of variance and the control of bias errors are two aspects of the sampling design which should be carefully considered in redesigning the sample. In the past, a number of summaries have been prepared on the observed data only. This ignores the probability sampling that was used and treats the data as if all hospitals had been selected with equal probability. It is clear that reporting of data from the NEISS should be based on the national estimates rather than on observed case frequencies, a change which is currently being partially implemented by the CPSC.

The summaries of the data will continue to evolve and improve. However, it is important to be able to investigate trends in the data. Consequently, whenever a new summary or index is instituted, it should also be calculated on previous years' data so that comparisons over time can be made. This is particularly important to the evaluation of standards or educational efforts of the CPSC. Care should be taken in reporting indices which are based on an arbitrary scale to ensure that the indices are used only for ranking. To do this, one should make certain that the scale is such that it cannot be confused with the estimated total number of injuries or some other interpretable figure.

C. The Hazard Index

The hazard index has been widely publicized by the media and given a status which it does not deserve. In particular, products on it have been viewed as consumer hazards, a title which they may not deserve. Excess attention has been given to the product at the top of the list or to the few products at the top of the list. The CPSC cannot control what is published. However, after a product has been high on the hazard index and has consequently had a number of in-depth studies done to determine the actual role in the injuries, a publication of the results and conclusions by the CPSC would be useful. In particular, it may find that the product ranks high on the list solely due to extensive use of the product and that no undue risks are associated with the product. This finding would be reflected in the assignment of new in-depth studies. Since the hazard index currently does not incorporate any exposure or risk data, the shift in emphasis from one group of products to another might be important both to the manufacturer and to the CPSC. The manufacturers would like the public to know that their product ranks high only because it is popular, not because it is hazardous, and it is to the CPSC's advantage to be able to answer critics who claim inaction on the part of the commission because the same product ranks high or at the top for several years.

The injury categories and the severity categories attached to the injuries are subjective. As a result, many people will not totally agree with them. Development of an adequate injury severity scale is far in the future, if ever, but it might be more useful if the scale used by the CPSC could be coordinated with the trauma scales developed by the American Medical Association such as the Abbreviated Injury Scale or the Comprehensive Injury Scale. Coordination of these scales would make the data acceptable to more professionals in the medical areas. The severity weights attached to the categories are quite arbitrary, and the orderings of the products on the resultant scale depend quite heavily on the particular numerical scale used. Consequently, the resultant index is useful only for ranking the products and a consensus should be sought for the numerical scale to be used.

The age adjustments used by the CPSC (really age weightings rather than adjustments) are highly controversial. It is a very subjective judgment that collectively all persons over 15 are only as important as those under 15 even though there are 2.5 times as many of them. Yet this is the rationale behind the current age adjustment. It seems too simplistic to expect a single age weighting to make all products comparable. A ranking of products within age groupings seems preferable. If an age weighting is to be used, account should be taken of susceptibility by age both to accidents, and to injury if an accident occurs. Since this will be highly debatable, such a weighting should be based on empirical data to the extent possible. Since only the relative magnitudes of the different ages are important--that is, only the shape of the curve--this is a reasonably straightforward procedure.

Although the calculation of the hazard index should and will continue to evolve, when a change is made, the new index should be computed for earlier data so that trends over time can be observed. A measure of risk or hazard based on exposure should be developed. This will be difficult, and will probably need to be somewhat product-specific, but should be incorporated as soon as possible, since this may be expected to change the relative ranking of the products substantially. At least until the hazard index can incorporate exposure data and be based on risk, the current priorities for investigation of products as measured by the number of in-depth investigations sought should be reported. Conclusions resulting from the in-depth investigations should also be reported by the Commission. The hazard index should be so scaled as to preclude misinterpretation of the index value as the number of injuries observed or projected or some other interpretable figure.

For the products studied in this analysis, when measured on a consistent basis over time, little change in the injury association over the three-year period is exhibited. For bicycles, there is a very strong seasonal component in the estimated number of bicycle-associated injuries, together with a consistent linear increase over time, which is of much smaller magnitude. The trend over time of associated injuries must be established in order to estimate the effect of a standard or other action by the Commission on the accidental injuries associated with any product.

IX. REFERENCES

Blake, R., Jr. "Pedestrian and Bicycle Accidents". State of Idaho Department of Highways Planning of Traffic Division, September, 1973.

Brezina, E. and Kramer, M. "An Investigation of Rider, Bicycle and Environmental Variables in Urban Bicycle Collisions". Ontario Department of Transport, Technicla Bulletin SE-70-01, October, 1970.

California Highway Patrol Analysis Section. "California Bicycle Accidents - Their Relationship to Enforcement and Equipment Factors". November, 1974.

Chlapecka, T. W., et al. "Bicycle Accidents and Usage Among Elementary School Children in the United States". National Safety Council, March, 1975.

Dougherty, N. and Lawrence, W. "Bicycle Transportation". Office of Planning and Evaluation, United States Environmental Protection Agency, December, 1974.

Guichon, D. M. P. and Myles, S. T. "Bicycle Injuries: One Year Sample in Calgary". The Journal of Trauma, Vol. 15, No. 6, 1975.

MAUDEP. "Proceedings of the Seminar on Planning, Design, and Implementation of Bicycle/Pedestrian Facilities, San Diego, 1974". The Institute of Transportation and Traffic Engineering, University of California at Berkeley, May, 1975.

National Safety Council. "Accidents Facts". 1970, 1972, 1974.

National Transportation Board. "Special Study: Bicycle Use as a Highway Safety Problem". Report No. NTSB-HSS-72-1, April 5, 1972.

Pascarella, E. A., et al. "Characteristics of Youthful Bicycle Riders in an Urban Community and Events Accruing to Operation of their Vehicles". University of North Carolina Highway Safety Research Center, June, 1971.

Popsh, L. N. and Lytel, R. B. "A Study of Bicycle-Motor Vehicle Accidents". Public Works Department, City of Santa Barbara, June, 1973.

Rice, R. S. and Roland, R. D., Jr. "An Evaluation of the Performance and Handling Qualities of Bicycles". Cornell Aeronautical Laboratory, Inc., Technical Report No. VJ-2888-K, April, 1970.

United States Bureau of the Census. 1970 Census of the Population: Number of Inhabitants, United States Summary. Vol. 1, No. 1 A, December, 1971.

United States Department of Transportation. "Bicycles USA." Proceedings, Conference, May 7-8, 1973.

Vilardo, F. J. and Anderson, J. H. "Bicycle Accidents to School Aged Children". National Safety Council, Report No. 169, September, 1969.

Vilardo, F. J., et al. "An Investigation into Bicycle Usage". National Safety Council, Report No. 268, September, 1968.

Waller, J. A. "Bicycle Ownership, Use and Injury Patterns Among Elementary School Children in Chittenden County, Vermont". August, 1970.

Waller, P. A. and Reinfurt, D. W. "Bicycles: An Analysis of Accidents in North Carolina". University of North Carolina Highway Safety Research Center, February, 1969.

Williams, A. F. "Factors in the Initiation of Bicycle-Motor Vehicle Collisions". The Insurance Institute for Highway Safety, December, 1974.

APPENDIX A

NEISS SYSTEM A; ESTIMATOR EQUATIONS

A report to the

Division of Consumer Product Safety

Bureau of Product Safety

Food and Drug Administration

PHS, DHEW

by

Thomas D. Downs

Associate Professor of Biometry

The University of Texas at Houston

School of Public Health

P.O. Box 20186

Houston, Texas 77025

November, 1971

TABLE OF CONTENTS

NEISS SYSTEM A ESTIMATOR EQUATIONS

1. INTRODUCTION

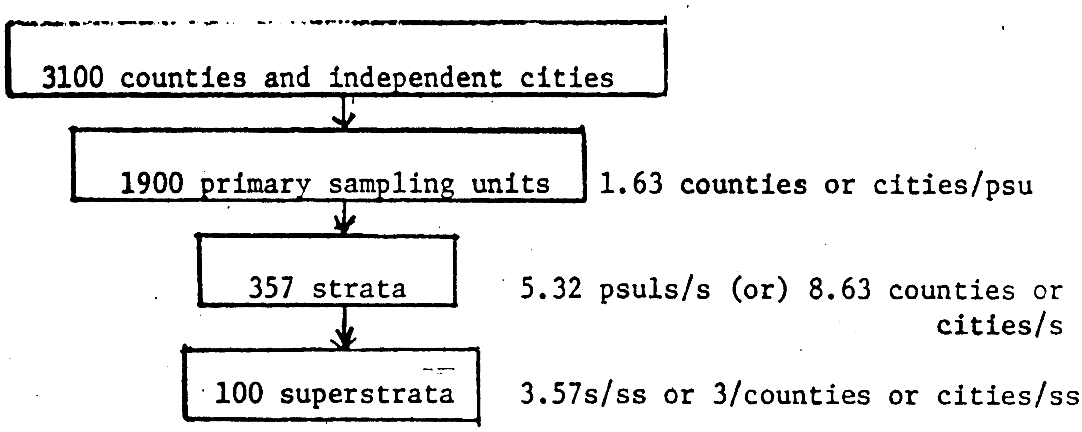
Purpose. The purpose of this document is to portray the mathematical equations for national estimates of consumer product injuries, based on NEISS System A data.

The primary responsibility fo the Division of Consumer Product Safety is to protect the American consumer against unduly hazardous or potentially hazardous products used in and around the home or in recreational areas. To fulfill this responsibility the Division of Consumer Product Safety must first determine how many of such injuries occur, the severities of these injuries, and the types of products involved. The estimator equations permit national estimates of various items of interest, such as the number of injuries of a specific nature. These national estimates, in turn, may be used to provide program direction, such as:

1. Identify the magnitudes of these injury problems.
2. Identify specific consumer products directly related to such injuries.
3. Suggest the need for standards, safety actions, or remedial actions.
4. Provide necessary data for evaluating the effectivenss of those items in 3, where action has been taken.
5. Assist in planning future improvements to the NEISS System A.

Background. The Division of Consumer Product Safety is currently implementing a National Electronic Injury Surveillance System (NEISS). The data collection of injuries by means of querying a probability sample of hospital emergency units is designated as System A within NEISS. The sampling plan used is similar to that for the Health Interview Survey (HIS), which itself is similar to that of the Current Population Survey (CPS)[1].

For sampling purposes the CPS constructed the following hierarchical structure for the United States: There are some 3100 counties and independent cities in the U.S. These were combined in various ways to form 1900 contiguous units called primary sampling units. These 1900 were subsequently combined to form 357 strata. Finally, the 357 strata were combined to form 100 superstrata. The sampling plan of the CPS (and also the HIS) is built around this hiearchical structure. A schematic representation of the structure is given below:



---

1 U.S. Bureau of the Census: The Current Population Survey-- A Report on Methodology. Technical Paper No. 7. Washington, D. C. U.S. Government Printing Office, 1963.

Of the 4906 hospitals in the sampling frame, 2606 appear in the 257 strata eliminated by Stage I, and the remaining 2300 appear in the 100 strata determined by Stage I. Subsequent sampling (Stage II) is confined to sampling from these 2300 hospitals. A schematic presentation of Stage I is given below:

100 superstrata
with 4906 hospitals

↓

100 strata (one from each superstratum)
with 2300 hospitals

<u>NEISS: Stage II</u>. Stage II consists of sampling, from the 2300 hospitals determined by Stage I, the actual individual hospitals to be included in the survey.

A useful method of increasing the precision of injury estimates is to stratify the hospitals in the sample frame into homogeneous-units called strata. Then at least two hospitals are sampled from each stratum. For NEISS System A stratification was done on a geographical and hospital-size basis. This process yielded 39 strata internally homogeneous with regard to geographical location and hospital size.

The geographical portion of the stratification was done by combining the 100 strata determined by Stage I into 13 <u>Blocks</u> of approximately equal human population. The Blocks are relatively homogeneous geographically. Each Block is completely contained in one of the four regions of the U.S. (Northeast, North Central, South, or West).

The stratification by hospital-size was done by using the number of emergency room visits (ERV's) in 1968 across hospitals. To this end, the

2. THE SAMPLING PLAN

The NEISS System A sampling frame consists of those hospitals in the U.S. which had emergency units in 1968, as determined by the American Hospital Association at that time. Long-term hospitals and federal penal hospitals are excluded from the frame. After these exclusions, there remain 4906 hospitals with emergency units, and it is from these 4906 that the sample of hospitals is to be drawn.

The sampling occurs in 2 stages. Stage I consists of selecting one stratum from each of the 100 superstrata. This yields 100 strata, and Stage II consists of sampling from within these 100 strata the actual individual hospitals to be included in the survey.

NEISS: Stage I. For each of the 100 superstrata, the stratum chosen to represent it by Stage I is selected at random on a population proportional basis. The probability of selecting a particular stratum within a superstratum is equal to

$$\frac{\text{human population* of the particular stratum}}{\text{human population* of the superstratum}}.$$

For example, if a superstratum consists of 3 strata whose populations are 1000, 2000, and 3000, then the probabilities that the first, second, or third of these would be selected to represent the superstratum would be 1/6, 2/6, and 3/6, respectively.

---

* Human population is based on the 1960 U.S. Census

The number of ERV's for each of the 2300 hospitals during 1968 was determined from existing records. When these were not reported the number of ERV's was estimated according to defined rules. Then the hospitals within each of the 13 Blocks were ranked by their determined number of 1968 ERV's. Next each Block was decomposed into Subblocks 1, 2, and 3 as follows:

For each Block, Subblock 1 for that Block was defined as the initial group of hospitals on the rank ordered listing accounting for 50% of the total number of ERV's within the Block. The group of hospitals accounting for the next 30% was denoted as Subblock 2. The group of smallest hospitals, accounting for the remaining 20%, was called Subblock 3. This process yielded the 39=13 x 3 strata.

As an example of Subblock determinations, suppose a Block had 10 hospitals within it, as determined by Stage I, and the rank ordering of these hospitals by number of ERV's was as follows:

Hospital	H_1	H_2	H_3	H_4	H_5	H_6	H_7	H_8	H_9	H_{10}	Total
No. of ERV's	10	8	7	6	5	4	4	3	2	1	50
Subblock		1			2			3			

Subblock 1 consists of the largest hospitals accounting for 25 ERV's (50% of 50), Subblock 2 accounts for 15 ERV's (30% of 50), and Subblock 3 accounts for 10 (20% of 50).

The actual distribution of the 2300 hospitals determined in Stage I, by Block and Subblock, is given below (Table 1).

TABLE 1

Distribution of the 2300 Hospitals determined in Stage I,

by Block and Subblock

Block	Subblock 1	Subblock 2	Subblock 3	Total
1.	28	55	145	228
2.	38	51	110	199
3.	21	29	95	145
4.	32	56	130	218
5.	18	30	89	137
6.	19	27	119	165
7.	21	30	112	163
8.	16	22	63	101
9.	5	16	108	129
10.	16	33	111	160
11.	16	46	145	207
12.	27	56	174	257
13.	29	47	115	191
Total	286	498	1516	2300

The 39 cells in Table 1 are the final strata used for NEISS System A. All estimates are based on these 39 strata.

It remains to determine the number of hospitals to be sampled within each stratum, and to specify the methods for selecting them. The sampling rates are, for each block, about 18% for Subblock 1, 7% for Subblock 2, and 2% for Subblock 3. Subblocks with larger hospitals are sampled at greater rates because the larger hospitals have more variation.

The total number of hospitals sampled is 119, and their distribution, by Block and Subblock, is given in Table 2.
The 5 hospitals shown in Table 2 for Subblock 1, Block 1, for instance, are randomly selected without replacement from the 28 hospitals for the same stratum in Table 1. Roughly speaking, the probabilities of selection are population proportional on a 1968 ERV basis.

TABLE 2

Distribution of the 119 Hospitals sampled in Stage II,

by Block and Subblock

Block	Subblock 1	Subblock 2	Subblock 3	Total
1.	5	4	3	12
2.	6	3	2	11
3.	4	2	2	8
4.	5	4	3	12
5.	3	2	2	7
6.	3	2	3	8
7.	4	2	2	8
8.	3	2	2	7
9.	2	2	2	6
10.	3	2	2	7
11.	3	3	3	9
12.	5	4	4	13
13.	5	3	3	11
Total	51	35	33	119

3. PROBABILITY CALCULATIONS

Data Required. Consider an arbitrary cell from the 39 cells in Table 2, and an arbitrary hospital, H_x, in that cell. A general expression will be derived for the probability, $P(H_x)$, that this hospital is included in the sample of 119 hospitals. This general expression may then be used, at least in theory, to calculate such probabilities for each of the 119 hospitals in the sample. To calculate $P(H_x)$ we require:

1) s=the 1960 human population of the CPS Stratum containing H_x.
2) S=the 1960 human population of the CPS superstratum containing H_x.
3) m=the number of hospitals (Table 1) in the cell containing H_x.
4) r=the number of hospitals sampled (Table 2) in the cell containing H_x.
5) A rank ordering of the m hospitals by number of ERV's:

Hospital:	$H_1, H_2, \ldots, H_m$.
No. of ERV's:	$n_1 \geq n_2 \geq \ldots \geq n_m$.

Items 30, 40 and 50 are common to every hospital in the same cell of Table 2, but items 1) and 2) are not necessarily so.

Probability Formulas. With the above notation, r hospitals are randomly selected, without replacement, from the m hospitals in the cell, these m hospitals being determined by Stage I. The probability that H_x is selected is the product of a) the probability that the CPS stratum containing H_x is selected in Stage I, and b) the probability that H_x is selected as one of the r hospitals in Stage II, given that the CPS stratum containing H_x was selected in Stage I.

The probability in a) is simply s/S. For b), H_x can be selected first, second, ..., or last. Let p_1 be the probability that H_x is selected first, p_2 the probability H_x is selected second, etc. to P_r being the probability H_x is selected last. The probability for b) is then $p_1+p_2+\ldots+P_r$, and thus

$$P(H_x) = \frac{s}{S}(p_1+p_2+\ldots+p_r). \qquad (0)$$

Let $n=n_1+n_2 \ldots+n_m$ be the total number of ERV's for the m hospitals in the cell. Then the probability that H_x is chosen first is $p_1=n_x/n$. The probability that H_x is chosen second is

$$P = \sum^{*} P(\text{n different hospital first})\ P(H_x \text{ second, given the first hospital}),$$

where the * indicates the summation is to be done over all hospitals except H_x, since H_x is to be chosen second and sampling is done without replacement. Thus

$$p_2 = \sum \frac{n_i}{n}\ \frac{n_x}{n-n_i},$$

the summation being over the (m-1) i's different from x. This last equation may be rewritten as

$$p_2 = \frac{n_x}{n}\sum^{*}\frac{n_i}{n-n_i}$$

Similarly, the probability that H_x is chosen third is

$$p_3 = \frac{n_x}{n}\sum^{*}\frac{n_i}{n-n_i}\ \frac{n_j}{n-n_i-n_j}$$

the summation being over all i and j (from 1 to m) for which $i\neq x$, $j\neq x$, and $i\neq j$. Thus there are (m-1)(m-2) terms to be summed, no two of the indices i,j,x being equal in any term.

According to Table 2 r is at most 6, so only p_1, p_2, ..., p_6 require mathematical expression. These formulas are:

$$p_1 = \frac{n_x}{n} \tag{1}$$

$$p_2 = \frac{n_x}{n} \sum_{i}^{*} \frac{n_i}{n-n_i} \tag{2}$$

$$p_3 = \frac{n_x}{n} \sum_{ij}^{*} \frac{n_i n_j}{(n-n_i)(n-n_i-n_j)} \tag{3}$$

$$p_4 = \frac{n_x}{n} \sum_{ijk}^{*} \frac{n_i n_j n_k}{(n-n_i)(n-n_i-n_j)(n-n_i-n_j-n_k)} \tag{4}$$

$$p_5 = \frac{n_x}{n} \sum_{ijk\ell}^{*} \frac{n_i n_j n_k n_\ell}{(n-n_i)(n-n_i-n_j)(n-n_i-n_j-n_k)(n-n_i-n_j-n_k-n_\ell)} \tag{5}$$

$$p_6 = \frac{n_x}{n} \sum_{ijk\ell t}^{*} \frac{n_i n_j n_k n_\ell n_t}{(n-n_i)(n-n_i-n_j)(n-n_i-n_j-n_k)(n-n_i-n_j-n_k-n_\ell)(n-n_i-n_j-n_k-n_\ell-n_t)} \tag{6}$$

each summation being over all possible permutations of the indices. Thus p_1 has 1 term, p_2 has m-1, p_3 has (m-1)(m-2), p_4 has (m-1)(m-2)(m-3), p_5 has (m-1)(m-2)(m-3)(m-4), and p_6 has (m-1)(m-2)(m-3)(m-4)(m-5) terms. Equations (2) through (6) above hold for every cell of Table 2, regardless of what r is. For any particular cell of Table 2 though, where r hospitals are sampled, only p_1 through p_r need be calculated, as per equation (0).

Example. Suppose for simplicity that there are m=4 hospitals in a certain cell, and r=3 of them are to be sampled. Suppose the data for the 4 hospitals are as follows:

	Hospital			
Data	H_1	H_2	H_3	H_4
s	3	4	5	6
S	10	20	30	40
n_i	4	3	2	1 , n=10

We shall find the probability, $P(H_x)$, that H_x is in the sample when x=4:

$$s/S = 6/40$$

$$p_1 = \frac{n_4}{n} = \frac{1}{10}$$

$$p_2 = \frac{n_4}{n}\left(\frac{n_1}{n-n_1} + \frac{n_2}{n-n_2} + \frac{n_3}{n-n_3}\right) = \frac{1}{10}\left(\frac{4}{6} + \frac{3}{7} + \frac{2}{8}\right) = \frac{113}{840} \sim 0.1345$$

$$p_3 = \frac{n_4}{n}\left[\frac{n_1}{n-n_1}\left(\frac{n_2}{n-n_1-n_2} + \frac{n_3}{n-n_1-n_3}\right) + \frac{n_2}{n-n_2}\left(\frac{n_1}{n-n_2-n_1} + \frac{n_3}{n-n_2-n_3}\right) + \right.$$

$$\left.\frac{n_3}{n-n_3}\left(\frac{n_1}{n-n_3-n_1} + \frac{n_2}{n-n_3-n_2}\right)\right]$$

$$= \frac{1}{10}\left[\frac{4}{6}\left(\frac{3}{3} + \frac{2}{4}\right) + \frac{3}{7}\left(\frac{4}{3} + \frac{2}{5}\right) + \frac{2}{8}\left(\frac{4}{4} + \frac{3}{5}\right)\right] \sim 0.2143$$

$$P(H_x) = \frac{s}{S}(p_1 + p_2 + p_3) = \frac{6}{40}(0.1000 + 0.1345 + 0.2143) \sim 0.0673$$

Computer Calculations. The number of terms summed in the calculation of p_r, for a hospital H_x in a cell of Table 1 with m hospitals, is (m-1)(m-2) ...(m-r+1). This is the number of permutations of m-1 things taken r-1 at a time, and is approximately m to the power r-1. This number can be reduced by a factor of (r-1)(r-2)...3.2.1 by expressing formulas (1) through (6) as summations over combinations rather than permutations. The appropriate formulas, (1') through (6'), are given below, and their derivations appear in the appendix. A computer program was written to calculate p_1 through p_r and this, together with a documentation for it, also appears in the appendix. The computer program was checked on a hypothetical cell of Table 2 where m=20, r=6, and the numbers of ERV;s for the 20 hospitals were in the ratios

1:2:3:4:...:18:19:20

A second check was done with m=20 and r=6 also, but with the hospital ERV's now in the ratios

21:22:23:24:...:38:39:40

The program was written in Fortran IV and ran on a Sigma Five digital computer. Total computation time for the two checks was slightly under one minute, or slightly under half a minute for each.

The biggest computational problems will occur in Block 2, Subblock 1 where m is about 40 and r is 6, and in Block 12, Subblock 3 were m is about 200 and r is 4. The first of these requires about 40^5, or 2^5 times as many terms as the check run. Since 2^5=32 the probabilities for this cell can be calculated in about 15 minutes on a computer comparable to a Sigma Five. The second requires about 200^3, or about twice as many terms as the check run, and should only take about one minute. All required probabilities can be calculated in an hour or two.

The combinatorial formulae for p_1 through p_6 are:

$$p_1 = n_x/n \tag{1'}$$

$$p_2 = p_1 C_2 \tag{2'}$$

$$p_3 = p_1[C_3-\binom{m-2}{1}C_2] \tag{3'}$$

$$p_4 = p_1[C_4-\binom{m-3}{1}C_3+\binom{m-2}{2}C_2] \tag{4'}$$

$$p_5 = p_1[C_5-\binom{m-4}{1}C_4+\binom{m-3}{2}C_3-\binom{m-2}{3}C_2] \tag{5'}$$

$$p_6 = p_1[C_6-\binom{m-5}{1}C_5+\binom{m-4}{2}C_4-\binom{m-3}{3}C_3+\binom{m-2}{4}C_2] \tag{6'}$$

where

$$\binom{a}{b} = \frac{a!}{b!\,(a-b)!}$$

and

$$C_{k+1} = \sum_{(i_1 \ldots i_k)}^{*} \frac{n_{i_1} + \ldots + n_{i_k}}{n - n_{i_1} - \ldots - n_{i_k}} \tag{7}$$

the sum being over all $\binom{m-1}{k}$ combinations of indices $i_1, \ldots, i_k$ taken from the first m integers but excluding x.

4. ESTIMATION

Notation. Let some characteristic of interest, C, be given. For example, C might be

> C = a male between 40 and 50 years of age going to a hospital emergency unit to obtain treatment for an accidental injury associated with a power lawn mower.

It is assumed that each ERV included in the sample can be classified as either having, or not having, the characteristic C. Let

> Y = total number of ERV's in the U.S. over a one year period which have the characteristic C.

Y is the quantity to be estimated. Let

$h=1,\ldots,13$ denote Block subscript (Table 1);

$i=1,2,3,$ denote Subblock subscript (Table 1);

$j=1,\ldots r_{hi}$ denote sampled hospital within the $(hi)^{th}$ cell of Table 1, where

r_{hi} = number of hospitals sampled in the $(hi)^{th}$ cell (Table 2);

n_{hij} = number of ERV;s within the j^{th} hospital of the $(hi)^{th}$ cell for the one year period in question;

w_{hij} = reciprocal of the probability, $(p_1+\ldots+p_{r_{hi}})\ /S$, that the j^{th} hospital of the $(hi)^{th}$ cell will be included in the sample;

q_{hij} = reciprocal $(1/p_1)$ of the probability, p_1, that this hospital was selected first for the $(hi)^{th}$ cell in Stage II.

Finally, let

$$Y_{hijk} \begin{cases} 1, & \text{if the } k^{th} \text{ ERV for the } j^{th} \text{ hospital of the } (hi)^{th} \text{ cell has the characteristic C.} \\ 0, & \text{if not.} \end{cases}$$

The Estimator Equations. The number of ERV's having the characteristic C for the j^{th} hospital of the $(hi)^{th}$ cell is

$$\hat{Y}_{hij} = \sum_{k=1}^{n_{hij}} Y_{hijk} \,,$$

and the estimated number of ERV's with C for the $(hi)^{th}$ cell is a weighted sum of these across the r_{hi} hospitals sampled in the cell, the weights being the reciprocals, w_{hij}:

$$\hat{Y}_{hi} = \sum_{j=1}^{r_{hi}} W_{hij} \hat{Y}_{hij}$$

Finally, the estimate $\hat{Y}$ of Y is the sum of the $\hat{Y}_{hi}$ over all cells:

$$\hat{Y} = \sum_{h=1}^{13} \sum_{i=1}^{3} \hat{Y}_{hi}$$

Combining the last two equations, the estimate $\hat{Y}$ of the total number of ERV's in the U.S. over a one year period which have the characteristic C is given by

$$\hat{Y} = \sum_{h=1}^{13} \sum_{i=1}^{3} \sum_{j=1}^{r_{hi}} W_{hij} \hat{Y}_{hij} \,.$$

This is the equation for actual calculation of $\hat{Y}$. It is calculated from 2 sets of 119 numbers each. The first set, w_{hij}, consists of the reciprocals of the probabilities of selection for the hospitals. The second set, $\hat{Y}_{hij}$, consists of the total counts of ERV's having the characteristic C, by hospital.

Variance Estimation. Sampling of hospitals within a given cell is independent of sampling within other cells, so the variance of $\hat{Y}$ is the sum of the variances of the $\hat{Y}_{hi}$ across all 39 cells. The variance of $\hat{Y}_{hi}$ depends only on between hospital variation within the $(hi)^{th}$ cell. There is no within-hospital variation because sampling is complete within hospitals.

The variation in the set $q_{hij} \hat{Y}_{hij}$ may be expected to be less than the variation in the set $\hat{Y}_{hij}$, since if history repeats itself and the number of ERV's with C is proportional to the total 1968 ERV's for each hospital, the set $c_{hij} \hat{Y}_{hij}$ will have zero variance. Thus

$$var(\hat{Y}_{hi}) = var \sum_{j=1}^{r_{hi}} \left(\frac{W_{hij}}{q_{hij}}\right)(q_{hij} \hat{Y}_{hij}) = \sum_{j=1}^{r_{hi}} \left(\frac{W_{hij}}{q_{hij}}\right)^2 var\{q_{hij} \hat{Y}_{hij}\},$$

Now let $(V_{hi})^2$ be defined by

$$(V_{hi})^2 = \sum_{j=1}^{r_{hi}} \left(\frac{W_{hij}}{q_{hij}}\right)^2 \tag{9}$$

The variance of $q_{hij} \hat{Y}_{hij}$ is estimated by

$$var(q_{hij} \hat{Y}_{hij}) = \frac{1}{r_{hi}-1}\left[\sum_{j=1}^{r_{hi}} (q_{hij} \hat{Y}_{hij})^2 - \frac{\left(\sum_{j=1}^{r_{hi}} q_{hij} \hat{Y}_{hij}\right)^2}{r_{hi}}\right]$$

for each j. Combining the above equations, the variance of $\hat{Y}$ is

$$var(\hat{Y}) = \sum_{h=1}^{13} \sum_{i=1}^{3} \frac{(V_{hi})^2}{r_{hi}-1}\left[\sum_{j=1}^{r_{hi}} (q_{hij} \hat{Y}_{hij})^2 - \frac{\left(\sum_{j=1}^{r_{hi}} q_{hij} \hat{Y}_{hij}\right)^2}{r_{hi}}\right] \tag{10}$$

Then, for example, 95% confidence limits for Y are given by

$$\hat{Y} \pm 1.96 \sqrt{var(\hat{Y})}. \tag{11}$$

The variance estimate is based on 3 sets of numbers:

39 of the $(V_{hi})^2$,

119 of the $q_{hij} \hat{Y}_{hij}$,

39 of the r_{hi}.

The $(V_{hi})^2$ are constant whatever the characteristic C - only the $\hat{Y}_{hij}$ change when C changes.

Regional estimates, $\hat{Y}$ and var $(\hat{Y})$, can be obtained from equations (8) and (10) by summing only over the Blocks h contained in the region.

5. APPENDIX

Derivation of Equations (1') through (6'). For what follows let

$$\left.\begin{array}{l} N_i = n - n_i , \; N_{ij} = n - n_i - n_j , \text{ etc.,} \\ n_{ij} = n_i + n_j , \; n_{ijk} = n_i + n_j + n_k , \text{ etc.,} \end{array}\right\} \quad (A1)$$

and let

$$T_2 = \sum_i^{*} \frac{n_i}{N_i} , \quad T_3 = \sum_{ij}^{*} \frac{n_i n_j}{N_i N_{ij}} , \quad T_4 = \sum_{ijk}^{*} \frac{n_i n_j n_k}{N_j N_{ij} N_{ijk}} , \text{ etc.,} \quad (A2)$$

where the sums are over all permutations of the indices excluding x.

Then from equations (A1), and (1) through (6), $p_1 = n_x/n$ and

$$p_2 = \frac{n_x}{n} T_2 , \quad p_3 = \frac{n_x}{n} T_3 , \quad p_4 \; \frac{n_x}{n} T_4 , \text{ etc.} \quad (A3)$$

Next define

$$S_2 = \sum_i^{*} \frac{n_i}{N_i} , \quad S_3 = \sum_{ij}^{*} \frac{n_i}{N_{ij}} , \quad S_4 = \sum_{ijk}^{*} \frac{n_i}{N_{ijk}} , \text{ etc.,} \quad (A4)$$

Where the sums here are also over all permutations of the indices excluding x. Finally, define

$$C_2 = \sum_i^{\circ} \frac{n_i}{N_i} , \quad C_3 = \sum_{ij}^{\circ} \frac{n_{ij}}{N_{ij}} , \quad C_4 = \sum_{ijk}^{\circ} \frac{n_{ijk}}{N_{ijk}} , \text{ etc.,} \quad (A5)$$

where now the sums are over all combinations of the indices excluding x, as denoted by the circles above the summation signs. Denote by P_b^a the number of permutations of a things taken b at a time:

$$P_b^a = a!/(a-b)! .$$

The derivation of (1') - (6') consists of showing that

$$\left.\begin{array}{l} 0!\,T_2 = S_2 \\ 1!\,T_3 = S_3 - P_1^{m-2} S_2 \\ 2!\,T_4 = S_4 - 2P_1^{m-3} S_3 + P_2^{m-2} S_2 \\ 3!\,T_5 = S_5 - 3P_1^{m-4} S_4 + 3P_2^{m-3} S_3 - P_3^{m-2} S_2 \\ 4!\,T_6 = S_6 - 4P_1^{m-5} S_5 + 6P_2^{m-4} S_4 - 4P_3^{m-3} S_3 + P_4^{m-2} S_2 \end{array}\right\} \quad (A6)$$

and showing that

$S_k=(K-2)!C_k$, K=2, ...,6. (A7)

Substituting (A&) in equations (A6), multiplying each side of the resulting equations by n_x/n, and using (A3) yields the desired equations (1') through (6').

The derivation of equations (A6) will be limited to showing that $T_3=S_3-P_1^{m-2}S_2$. The others are derived similarly. Thus, using the fact that $\frac{a}{b(b-a)}=\frac{1}{b-a}-\frac{1}{b}$, we get

$$T_3=\sum_{ij}^{*}\frac{n_i n_j}{N_i N_{ij}}=\sum_{ij}^{*}n_i\left(\frac{1}{N_{ij}}-\frac{1}{N_i}\right)=\sum_{ij}^{*}\frac{n_i}{N_{ij}}-\sum_{ij}^{*}\frac{n_i}{N_i}$$

$$=S_2-(m-2)\sum_{i}^{*}\frac{n_i}{N_i}=S_3-(m-2)S_2=S_3-P_1^{m-2}S_2.$$

To show (A7) with K=4, for example, note that by symmetry

$$S_4=\sum_{ijk}^{*}\frac{n_i}{N_{ijk}}=\sum_{ijk}^{*}\frac{n_j}{N_{ijk}}=\sum_{ijk}^{*}\frac{n_k}{N_{ijk}}$$

so that

$$3S_4=\sum_{ijk}^{*}\frac{n_{ijk}}{N_{ijk}},$$

and the sum is over all permutations of i,j,k excluding x. But for fixed i,j,k the quantity n_{ijk}/N_{ijk} is invariant under the 3 permutations of i,j,k. Hence

$$3S_4=\sum_{ijk}^{*}\frac{n_{ijk}}{N_{ijk}}=3!\sum_{ijk}^{o}\frac{n_{ijk}}{N_{ijk}}=3!\,C_4,$$

so that $S_4=(4-2)!C_4$. The general case is similarly shown. This completes the derivation of equations (1') through (6').

APPENDIX B

VARIANCE ESTIMATES

Let $\hat{Y}$ denote the estimate of the national yearly total of bicycle-associated injuries. Because of the stratification the estimate of variance of $\hat{Y}$ is given by

$$\hat{V}(\hat{Y}) = \sum_{i=1}^{3} \sum_{h=1}^{13} \hat{V}(\hat{Y}_{hi}),$$

where

$\hat{Y}_{hi}$ is the estimate from the $(hi)^{th}$ stratum. That is,

$$\hat{Y}_{hi} = \sum_{j=1}^{r_{hi}} W_{hij}\hat{Y}_{hij},$$

where r_{hi} is the number of hospitals sampled in the $(hi)^{th}$ stratum and $\hat{Y}_{hij}$ is the number of bicycle-associated injuries observed in the j^{th} hospital in the $(hi)^{th}$ stratum.

Then the estimated variance of $\hat{Y}_{hi}$ is given by

$$\hat{V}(\hat{Y}_{hi}) = \frac{1}{r_{hi}-1}\left[\sum_{j=1}^{r_{hi}} (W_{hij}\hat{Y}_{hij})^2 - \frac{\{\sum_{j=1}^{r_{hi}} (W_{hij}\hat{Y}_{hij})\}^2}{r_{hi}}\right]$$

This may, of course, be written as

$$\hat{V}(\hat{Y}_{hi}) = \frac{1}{r_{hi}-1}\left[\sum_{j=1}^{r_{hi}} \left(\frac{W_{hij}}{q_{hij}}\right)^2 (q_{hij}\hat{Y}_{hij})^2 - \frac{\left\{\sum_{j=1}^{r_{hi}} \left(\frac{W_{hij}}{q_{hij}}\right) q_{hij}\hat{Y}_{hij}\right\}^2}{r_{hi}}\right]$$

since the q_{hij} cancel out. There seems to be no advantage to the introduction of the q_{hij}, however.

The estimate of the variance within each stratum may be found simply by taking the observed number of injuries (of a particular type - e.g., bicycle-associated) in each hospital, multiplying by the hospital weight, and then calculating the usual sample variance among the resulting r_{hi} numbers in each stratum. Aside from a number of obvious typographical errors in the original variance equations (reproduced in Appendix A), there is a basic error in the suggested computations. The estimated variance of $\hat{Y}_{hi}$ is found by calculating a sample variance among the numbers $\{W_{hi}, \hat{Y}_{hij}\}$, $j=1,\ldots r_{hi}$ in each stratum, not as

$$\text{Var}\left\{\sum_{j=1}^{r_{hi}} \left(\frac{W_{hij}}{q_{hij}}\right) q_{hij}\hat{Y}_{hij}\right\} .$$

The latter form would be appropriate if there were also a within-hospital component so that the $\hat{Y}_{hij}$ would themselves be subject to sampling variation. However, since a census of emergency room visits for in-scope products is obtained for each hospital, this term is not present.

APPENDIX C

ORIGINAL NEISS HOSPITALS

Block-Subblock	Hospital Number	Number Sampled	Index Number	Sequence Number	Hospital Name	Prob. of Selection 1st Stage	Prob. of Selection 2nd Stage	Combined Prob.
1-1		(5)						
	1-14-36-001		3	5	Harlem Hospital Center, New York City	1.00000000	.27856216	.27856216
	1-14-36-005		19	10	Bronx-Lebanon, New York City	1.00000000	.13278939	.13278939
	1-14-36-011		23	15	Brookdale, Brooklyn, New York	1.00000000	.12063632	.12063632
	1-14-36-013		24	20	Nassau Co. Medical, East Meadow, N.Y.	1.00000000	.10751031	.10751031
	1-14-36-002		28	25	Jewish Hospital, New York City	1.00000000	.09129410	.09129410
1-2		(4)						
	1-14-36-009		35	30	Beekman-Downtown, New York City	1.00000000	.06295267	.06295267
	1-14-34-007		44	35	Greater Paterson, New Jersey	1.00000000	.05810666	.05810666
	1-14-36-014		46	40	White Plains, New York	1.00000000	.05674560	.05674560
	1-14-34-006		49	45	Barnert Memorial, Paterson, New Jersey	1.00000000	.05537401	.05537401
1-3		(3)						
	1-14-36-010		25	50	French and Polyclinic, New York City	1.00000000	.04568325	.04568325
	1-14-36-003		33	55	St. John's, Riverside, New York City	1.00000000	.03823548	.03823548
	1-14-36-012		42	60	Peekskill Community, New York	1.00000000	.03188165	.03188165
2-1		(6)						
	1-14-36-016		6	65	Strong Memorial, Rochester, New York	.33755760	.21964890	.07414416
	1-14-36-008		18	70	Rochester General, New York	.33755760	.12975665	.04380034
	1-15-42-006		23	75	Lower Bucks, Bristol, Pennsylvania	1.00000000	.11530621	.11530621
	1-15-42-009		25	80	Mercy Catholic, Philadelphia, Pa.	1.00000000	.11318516	.11318516
	1-15-42-001		31	85	University of Pa., Philadelphia, Pa.	1.00000000	.10234419	.10234419
	1-14-34-008		38	90	Burlington Co. Mem., Mt. Holly, N.J.	1.00000000	.09456243	.09456243
2-2		(3)						
	1-03-25-001		2	95	Quincy City, Massachusetts	1.00000000	.07311317	.07311317
	1-15-42-004		24	100	Bryn Mawr, Pennsylvania	1.00000000	.06128795	.06128795
	1-15-42-003		44	105	Westmoreland, Greensburg, Pennsylvania	1.00000000	.04764211	.04764211
2-3		(2)						
	1-15-42-002		37	110	St. John's General, Pittsburgh, Pa.	1.00000000	.02336341	.02336341
	1-15-42-008		49	115	Eye and Ear of Pittsburgh, Pennsylvania	1.00000000	.01897735	.01897735

ORIGINAL NEISS HOSPITALS

Block-Subblock	Hospital Number	Number Sampled	Index Number	Sequence Number	Hospital Name	Prob. of Selection 1st Stage	Prob. of Selection 2nd Stage	Combined Prob.
3-1		(4)						
	1-14-36-015		3	120	Ed Meyer, Buffalo, New York	1.00000000	.25317467	.25317467
	1-03-25-004		4	125	Wesson Memorial, Springfield, Mass.	.30144745	.25169802	.07587373
	1-14-36-017		9	130	Mercy, Buffalo, New York	1.00000000	.17682915	.17682915
	1-03-25-002		20	135	Holyoke, Massachusetts	.30144745	.12837337	.03869782
3-2		(2)						
	1-15-42-007		1	140	Robert Packer, Sayre, Pennsylvania	.25066430	.09830607	.02464182
	1-14-36-004		3	145	Emergency, Buffalo, New York	1.00000000	.08642501	.08642501
3-3		(2)						
	1-14-36-007		36	150	Community, Monticello, New York	.18518519	.02313201	.00428371
	1-15-42-005		40	155	Gnaden-Huetten, Lehighton, Pennsylvania	.19176319	.02025433	.00388403
4-1		(5)						
	1-05-17-002		1	160	Cook General, Chicago, Illinois	1.00000000	.66483671	.66483671
	1-10-29-006		5	165	St. Louis City, Missouri	1.00000000	.27590127	.27590127
	1-06-26-001		11	170	Pontiac General, Michigan	1.00000000	.14726114	.14726114
	1-05-17-011		22	175	Naval Hospital, Great Lakes, Illinois	1.00000000	.09344789	.09344789
	1-10-29-004		27	180	Lutheran, St. Louis, Missouri	1.00000000	.08360040	.08360040
4-2		(4)						
	1-05-17-003		1	185	MacNeal, Berwyn, Illinois	1.00000000	.09583850	.09583850
	1-05-17-013		13	190	Silver Cross, Joliet, Illinois	1.00000000	.08683976	.08683976
	1-05-17-008		39	195	Roseland Community, Chicago, Illinois	1.00000000	.06204244	.06204244
	1-05-17-012		55	200	Holy Family, Des Plaines, Illinois	1.00000000	.04806911	.04806911
4-[illegible]		(3)						
	[illegible]		6	205	Alton Memorial, Illinois	1.00000000	.05031131	.05031131
	[illegible]		8	210	Louis Weiss, Chicago, Illinois	1.00000000	.04926746	.04926746
	[illegible]		[illegible]	215	Copley Memorial, Aurora, Illinois	1.00000000	.03437278	.03437278

REPLACEMENT NEISS HOSPITALS

Block-Subblock	Hospital Number	Number Sampled	Index Number	Sequence Number	Hospital Name	Prob. of Selection 1st Stage	Prob. of Selection 2nd Stage	Combined Prob.	Replacement Date
3-1		(4)							
	1-14-36-006		16	135	St. Joseph's, Syracuse, New York	.40869565	.15325321	.06263392	08-19-74
4-2	1-06-26-004	(4)	34	185	St. Joseph's Hospital, Mt. Clemons, Mich.	1.00000000	.06432171	.06432171	12-05-74

ORIGINAL NEISS HOSPITALS

Block-Subblock	Hospital Number	Number Sampled	Index Number	Sequence Number	Hospital Name	Prob. of Selection 1st Stage	Prob. of Selection 2nd Stage	Combined Prob.
5-1		(3)						
	1-06-39-004		8	220	Mercy, Hamilton, Ohio	.18855391	.16860645	.03179141
	1-06-26-002		14	225	Borgess, Kalamazoo, Michigan	.18855391	.13645593	.02572930
	1-15-42-010		17	230	Greenville, Pennsylvania	.33211679	.10007309	.03323595
5-2		(2)						
	1-10-19-001		6	235	St. Joseph's, Keokuk, Iowa	.22684211	.09199518	.02086838
	1-06-26-005		24	240	Community, Watervliet, Michigan	.17264344	.04970651	.00858150
5-3		(2)						
	1-12-55-004		31	245	St. Mary's, Sparta, Wisconsin	.25066525	.02647013	.00663514
	1-12-55-002		48	250	Apple River Valley, Amery, Wisconsin	.25066525	.01891101	.00474033
6-1		(3)						
	1-12-27-003		4	255	St. Paul-Ramsey, Minnesota	1.00000000	.20429602	.20429602
	1-06-39-006		10	260	Parma Community, Ohio	1.00000000	.13216427	.13216427
	1-06-39-005		19	265	Euclid General, Ohio	1.00000000	.10248067	.10248067
6-2		(2)						
	1-06-39-001		1	270	St. Vincent Charity, Cleveland, Ohio	1.00000000	.10669947	.10669947
	1-06-39-003		21	275	Forest City, Cleveland, Ohio	1.00000000	.05394502	.05394502
6-3		(3)						
	1-10-19-007		3	280	Bethesda General, Fort Dodge, Iowa	.17318436	.09190291	.01591615
	1-12-27-001		48	285	Sanford Memorial, Farmington, Minnesota	1.00000000	.01964176	.01964176
	1-08-38-002		51	290	Unity, Grafton, North Dakota	.23100490	.01913154	.00441948

REPLACEMENT NEISS HOSPITALS

Block-Subblock	Hospital Number	Number Sampled	Index Number	Sequence Number	Hospital Name	Prob. of Selection 1st Stage	Prob. of Selection 2nd Stage	Combined Prob.	Replacement Date
5-2		(2)							
	1-10-29-007		7	240	Levering Hospital, Hannibal, Missouri	.22684211	.09079751	.02059670	07-01-74

ORIGINAL NEISS HOSPITALS

Block-Subblock	Hospital Number	Number Sampled	Index Number	Sequence Number	Hospital Name	Prob. of Selection 1st Stage	Prob. of Selection 2nd Stage	Combined Prob.
7-1		(4)						
	1-12-55-007		1	295	Milwaukee County General, Wisconsin	1.00000000	.54690312	.54690312
	1-10-20-009		2	300	Munson, Fort Leavenworth, Kansas	.27881041	.33120798	.09234423
	1-12-27-006		14	305	St. Mary's, Rochester, Minn.	.24416667	.14437369	.03525124
	1-10-29-005		18	310	Independence Sanitarium, Missouri	1.00000000	.12778806	.12778806
7-2		(2)						
	1-07-40-003		12	315	Mission Hill, Shawnee, Oklahoma	1.00000000	.07183880	.07183880
	1-12-55-005		26	320	Milwaukee Children's, Wisconsin	1.00000000	.04590351	.04590351
7-3		(2)						
	1-06-39-002		51	325	Lodi Community, Ohio	.25877551	.01339671	.00346674
	1-10-19-002		64	330	Jane Lamb, Clinton, Iowa	.24416667	.00796214	.00194409
8-1		(3)						
	1-15-24-007		4	335	Sinai of Baltimore, Maryland	1.00000000	.20031907	.20031907
	1-15-24-005		7	340	Prince George's General, Cheverly, Md.	1.00000000	.17779817	.17779817
	1-15-24-004		16	345	St. Joseph's, Towson, Maryland	1.00000000	.12397382	.12397382
8-2		(2)						
	1-15-24-009		10	350	Lutheran of Maryland, Baltimore, Md.	1.00000000	.09641799	.09641799
	1-15-24-003		17	355	Washington Sanitarium, Takoma Park, Md.	1.00000000	.07626880	.07626880
8-3		(2)						
	1-01-13-016		3	360	Phoebe Putney, Albany, Georgia	.18750000	.08281986	.01552872
	1-01-37-011		7	365	Annie Penn, Reidsville, North Carolina	.23807132	.06604543	.01572352

ORIGINAL NEISS HOSPITALS

Block-Subblock	Hospital Number	Number Sampled	Index Number	Sequence Number	Hospital Name	Prob. of Selection 1st Stage	Prob. of Selection 2nd Stage	Combined Prob.
9-1		(2)						
	1-01-13-024		1	370	Fort Benning, Columbus, Georgia	.25529412	.69180013	.17661250
	1-01-12-004		3	375	Jackson Memorial, Miami, Florida	.60012837	.34007281	.20408734
9-2		(2)						
	1-01-13-017		1	380	Medical Center, Columbus, Georgia	.25529412	.20531893	.05241671
	1-01-12-002		11	385	Hialeah, Florida	.60012837	.08807371	.05285553
9-3		(2)						
	1-01-45-010		2	390	Tuomey, Sumter, South Carolina	.18299320	.08048627	.01472844
	1-01-12-003		7	395	Parkway, North Miami, Florida	.60012837	.06782073	.04070114
10-1		(3)						
	1-01-21-023		1	400	Fort Campbell, Kentucky	.30382596	.59076787	.17949061
	1-01-21-015		6	405	St. Joseph's Infirmary, Louisville, Ky.	.54634514	.15855087	.08662350
	1-15-54-008		14	410	St. Mary's, Huntington, West Virginia	.18279570	.10653072	.01947336
10-2		(2)						
	1-01-47-008		3	415	University of Tennessee, Knoxville, Tenn.	.26080794	.09309032	.02427869
	1-01-47-020		13	420	Fort Sanders, Knoxville, Tennessee	.26080794	.06719384	.01752469
10-3		(2)						
	1-01-12-018		37	425	Metropolitan General/Pineallas Park, Fla.	.57871064	.02255684	.01305388
	1-01-12-014		68	430	Campbellton-Graceville, Florida	.29731590	.01131428	.00336391

REPLACEMENT NEISS HOSPITALS

Block-Subblock	Hospital Number	Number Sampled	Index Number	Sequence Number	Hospital Name	Prob. of Selection 1st Stage	Prob. of Selection 2nd Stage	Combined Prob.	Replacemen Date
9-2	1-07-48005	(2)	5	380	Santa Rosa Med. Ctr., San Antonio, Texas	.42459827	.16320315	.06929578	03-01-75

ORIGINAL NEISS HOSPITALS

Block-Subblock	Hospital Number	Number Sampled	Index Number	Sequence Number	Hospital Name	Prob. of Selection 1st Stage	Prob. of Selection 2nd Stage	Combined Prob.
11-1		(3)						
	1-01-13-012		1	435	Grady Memorial, Atlanta, Georgia	1.00000000	.51255870	.51255870
	1-07-48-004		3	440	Parkland, Dallas, Texas	1.00000000	.38544591	.38544591
	1-01-13-001		8	445	Kennestone, Marietta, Georgia	1.00000000	.14328945	.14328945
11-2		(3)						
	1-15-51-006		4	450	Winchester, Virginia	.26858513	.10029843	.02693867
	1-07-48-002		21	455	Arlington, Texas	.54364326	.06156743	.03347072
	1-07-48-001		35	460	Glenview, Fort Worth, Texas	.54364326	.04550810	.02474017
11-3		(3)						
	1-01-28-009		7	465	Jeff Anderson, Meridian, Mississippi	.28783383	.05301058	.01525824
	1-01-21-019		54	470	Haggin, Harrodsburg, Kentucky	.32086614	.02439339	.00782701
	1-01-21-013		58	475	Woodford Memorial, Versailles, Kentucky	.32086614	.02303572	.00739138
12-1		(5)						
	1-11-06-013		6	480	Long Beach Naval, California	1.00000000	.24528607	.24528607
	1-11-06-001		11	485	Children's, Los Angeles, California	1.00000000	.17533615	.17533615
	1-11-06-012		12	490	Orange County Medical, Los Angeles, Calif.	1.00000000	.17495566	.17495566
	1-16-06-001		18	495	Brookside, San Pablo, California	1.00000000	.11391113	.11391113
	1-11-06-011		27	500	Pomona Valley, California	1.00000000	.09049050	.09049050
12-2		(4)						
	1-16-06-003		2	505	Mt. Zion, San Francisco, California	1.00000000	.10314446	.10314446
	1-11-06-002		6	510	Holy Cross, San Fernando, California	1.00000000	.09596459	.09596459
	1-16-06-004		7	515	Harold Chope, San Mateo, California	1.00000000	.09221632	.09221632
	1-11-06-010		44	520	Community, San Gabriel, California	1.00000000	.05422696	.05422696
12-3		(4)						
	1-11-06-015		10	525	Centinela Valley, Inglewood, California	1.00000000	.05843591	.05843591
	1-11-06-006		25	530	Memorial Southern Calif., Culver City, Calif.	1.00000000	.04901873	.04901873
	1-11-06-003		41	535	Garden Park, Anaheim, California	1.00000000	.03864885	.03864885
	1-17-[illegible]-006		48	540	St. Luke's, Bellingham, Washington	.21257143	.03629848	.00771602

REPLACEMENT NEISS HOSPITALS

Block-Subblock	Hospital Number	Number Sampled	Index Number	Sequence Number	Hospital Name	Prob. of Selection 1st Stage	Prob. of Selection 2nd Stage	Combined Prob.	Replacement Date
11-2	1-01-37-025	(3)	20	450	North Carolina Memorial, Chapel Hill, N.C.	.26858513	.06600475	.01772790	05-01-73
12-2	1-16-06-006	(4)	54	505	Pacific Med. Ctr., San Francisco, Calif.	1.00000000	.05027506	.05027506	03-01-7
	1-11-06-018		53	520	Memorial of Glendale, California	1.00000000	.05073955	.05073955	12-11-7
12-3	1-11-06-016	(4)	47	525	Burbank Community, California	1.00000000	.03722236	.03722236	05-01-7
	1-11-06-017		31	535	Garfield Hospital, Monterey Park, California	1.00000000	.04581956	.04581956	07-01-7

ORIGINAL NEISS HOSPITALS

Block-Subblock	Hospital Number	Number Sampled	Index Number	Sequence Number	Hospital Name	Prob. of Selection 1st Stage	Prob. of Selection 2nd Stage	Combined Prob.
13-1		(5)						
	1-16-06-002		5	545	Santa Clara Medical, San Jose, California	.40839695	.27653953	.11293790
	1-17-53-005		9	550	Harborview Medical, Seattle, Washington	1.00000000	.18926809	.18926809
	1-17-41-003		11	555	University of Oregon—North, Portland, Ore.	.62892119	.17695229	.11128905
	1-17-53-001		15	560	Holy Family, Spokane, Washington	.20981132	.13826733	.02901005
	1-08-49-001		19	565	McKay-Dee, Ogden, Utah	.30246020	.12818635	.03877127
13-2		(3)						
	1-08-30-003		25	570	Billings Deaconess, Montana	.30246020	.05932416	.01794320
	1-11-06-009		29	575	Simi Valley, California	.18158236	.05847409	.01061786
	1-11-06-007		45	580	Parkview, Riverside, California	.61832061	.04667459	.02885986
13-3		(3)						
	1-17-53-002		3	585	Swedish, Seattle, Washington	1.00000000	.06805927	.06805927
	1-17-53-004		4	590	General of Everett, Washington	1.00000000	.06758111	.06758111
	1-11-06-014		20	595	Dispensary, Marine Corps, 29 Palms, Calif.	.61832061	.04859038	.03004443

REPLACEMENT NEISS HOSPITALS

Block-Subblock	Hospital Number	Number Sampled	Index Number	Sequence Number	Hospital Name	Prob. of Selection 1st Stage	Prob. of Selection 2nd Stage	Combined Prob.	Replacement Date
13-2	1-08-08-004	(3)	23	570	St. Mary's-Corwin, Pueblo, Colorado	.19044586	.06052374	.01152650	03-01-7