15851

Report 03169-V

ACQUISITION OF INFORMATION ON EXPOSURE AND ON NON-FATAL CRASHES

Volume V - Executive Summary

Philip S. Carroll

Highway Safety Research Institute
The University of Michigan
Ann Arbor, Michigan 48104

May 12, 1971

Final Report

Prepared under contract FH-11-7293 for

National Highway Traffic Safety Administration
Department of Transportation
Washington, D.C. 20591

TECHNICAL REPORT STANDARD TITLE PAGE

1. Report No.	2. Government Accession No.	3. Recipient's Catalog No.
03169 - V		

4. Title and Subtitle	5. Report Date
Acquisition of Information on Exposure and on Non-Fatal Crashes Volume V - Executive Summary	May 12, 1971 6. Performing Organization Code
7. Author(s)	8. Performing Organization Report No.
Philip S. Carroll	
9. Performing Organization Name and Address	10. Work Unit No.
Highway Safety Research Institute The University of Michigan Ann Arbor, Michigan 48104	11. Contract or Grant No. FH-11-7293
12. Sponsoring Agency Name and Address	13. Type of Report and Period Covered Final Report June, 1969 to March, 1971.
National Highway Traffic Safety Admin. Washington, D.C. 20591	14. Sponsoring Agency Code FH-11

15. Supplementary Notes

The opinions, findings, and conclusions expressed in this publication are those of the authors and not necessarily those of the National Highway Traffic Safety Administration.

16. Abstract

This volume presents summaries of each of the three phases of the study.

In Phase I, dealing with exposure survey considerations, analyses were performed on classifications of exposure data and procedures for exposure surveys. A pilot survey of driving exposure (vehicle miles of travel) was performed and six variables were determined as best predictors of exposure: driver sex and age, vehicle type and model year, roadway type, and day vs. night. Considerations of cost and accuracy among various methods for exposure surveys led to a recommendation of annual, nationwide surveys using mailed questionnaries and one-day trip records by randomly sampled drivers.

In Phase II, dealing with accident data inaccuracies, data from a survey of drivers was compared with driver's offical records to determine under-reporting bias. Results may be used to make corrections in accident data. A technique for correcting accident injury data, based on hospital records, was considered.

In Phase III, detailed procedures were determined to implement a national exposure survey in 1972.

17. Key Words	18. Distribution Statement
	NTIS For public distribution

19. Security Classif. (of this report)	20. Security Classif. (of this page)	21. No. of Pages	22. Price
none	none	23	

Form DOT F 1700.7 (8-69)

Preface

This is Volume V (Executive Summary) of the final report on Contract FH-11-7293, "Acquisition of Information on Exposure and on Non-Fatal Crashes."

It presents basic concepts of driving exposure and accident-data inaccuracies, and summarizes the most important findings of the three phases of the study (See Volumes I, II, and III).

The total report is probably the first comprehensive attempt to analyze the needs for collection of exposure data on a large scale, and to show the relationship between exposure data and accident data in the calculation of accident rates. In this context, the recommendations are of fundamental importance in future evaluations of highway safety countermeasures.

A major purpose of this Executive Summary, therefore, is to provide a concise document for use in policy decisions regarding future implementation of exposure surveys, and programs to standardize the quality of accident data.

TABLE OF CONTENTS

LIST OF ILLUSTRATIONS

SECTION 1
INTRODUCTION

This volume presents a summary of the tasks and results in all three phases of Contract FH-11-7293, "Acquisition of Information on Exposure and on Non-Fatal Crashes." Its major purpose is to provide a concise document for use in policy decisions regarding future implementation of exposure surveys, and programs to standardize the quality of aggregate accident data.

OBJECTIVES

As stated in the contract, the objectives of the study are:

1. To forumlate a logical structure and methodology to aid in the orderly acquisition of exposure data.
2. To develop sampling techniques and procedures for obtaining mileage of travel on different classes of highways, with differing traffic characteristics, for significant driver-vehicle-combinations.
3. To develop procedures for obtaining reliable estimates of numbers and types of property damage and personal injury crashes and the associated damage and injuries.

A general aim in collecting reliable exposure data and reliable accident data is to permit the calculation of valid accident rates (accidents per vehicle mile) with accident occurrence frequency in the numerator and a corresponding exposure measure (mileage) in the denominator. Corresponding sets of accident data and exposure data are most useful when classified according to various driver-vehicle-road-environment combinations. In this way, meaningful comparisons may be made between different elements of the highway transportation system, thereby providing evaluations of the effectiveness of highway safety countermeasures. The results of this study are intended to create a basis for future evaluation programs.

ORGANIZATION

This volume is the Executive Summary of a five-volume final

report. The listing below indicates the volume numbers and titles, and where appropriate, the corresponding phases of the contract Statement of Work (Appendix A).

Volume I - Exposure Survey Considerations
(Phase I - Exposure Information)

Volume II - Accident Data Inaccuracies
(Phase II - Information on Non-Fatal Crashes)

Volume III - Procedures for an Exposure Survey
(Phase III - Driving Exposure Survey)

Volume IV - Appendices

Volume V - Executive Summary

The three main sections which follow summarize the results of volumes I, II, and III, respectively.

SECTION 2
EXPOSURE SURVEY CONSIDERATIONS

This section summarizes the results of Phase I (Exposure Information), as detailed in Volume I.

EXPOSURE CLASSIFICATIONS

The requirement of this task is to "determine the principal classes of drivers and vehicles and environments for which exposure measurements are needed." The classes may be characterized by variables such as driver age, vehicle type and road type, and combinations thereof. The classes are required to be "relatively homogeneous with respect to relevant exposure factors," amenable to sampling procedures, and "useful for studying the impact of safety countermeasures."

This task is the key task in the exposure study. Its results define the required content of future exposure data records, and they determine feasible alternatives for data collection in later tasks.

The approach was to perform two consecutive exposure surveys, and to analyze their data in terms of the variables which are best predictors of exposure. The first survey had a small sample (448), but it had a large number of potential "predictor" variables as candidates for the definition of unique exposure classes. The second survey had a very large sample size (8000), but it had a greatly reduced number of potential predictor variables. (Many of the variables of the first survey were eliminated because analysis showed they were not good predictors of exposure).

Both surveys were conducted by personal interviews of licensed drivers in licensing offices. The first was done in one office, and the second was done in 37 offices distributed throughout the country. Drivers were asked for estimates of mileage driven in the last 30 days, and for information on themselves, their vehicle, and their driving patterns.

Data analysis was performed by means of a computer program (AID - Automatic Interaction Detector) which divides data samples into smaller groups by picking the best predictor variable on the basis of minimum variance. Successive analyses of the smaller groups leads to a hierarchy of best predictor variables, as illustrated in Figure 1.

This figure identifies four variables which determine best splitting of sample groups: Drive on Job?, Driver Sex, Type of Vehicle Driven, and Percent Driving on Local Streets. The boxes indicate group size N, mean miles $\overline{Y}$, and variable levels which define one part of a two-way group split.

The first variable, Drive on Job?, does not correspond to any item of data on most accident reports. Therefore, it cannot be used to determine accident rates in the near future. However, Type of Vehicle Driven serves as a good substitute for Drive on Job? because of their strong correlation. But in the future, a Drive on Job variable should be considered for inclusion in accident reports.

Further AID runs using logarithm of miles and number of accidents as dependent variables produced three more recommended predictor variables: Driver Age, Model Year of Vehicle, and Percent Driving at Night.

Other variables not selected after the second survey were: number of vehicles driven, vehicle use, driver's knowledge of engine, education, income, car size and make, urbanization index, socio-economic index, area population, and percentages of driving on freeways, rural roads, and wet roads.

Figure 2 presents a chart of 26 unique exposure classes based on the six selected variables:

Vehicle Type (Passenger Cars and Small Trucks vs. Other Vehicle Types)
Driver Sex (Male vs. Female)
Road Type (Local Streets vs. Other Road Types)
Light Condition (Day vs. Night)

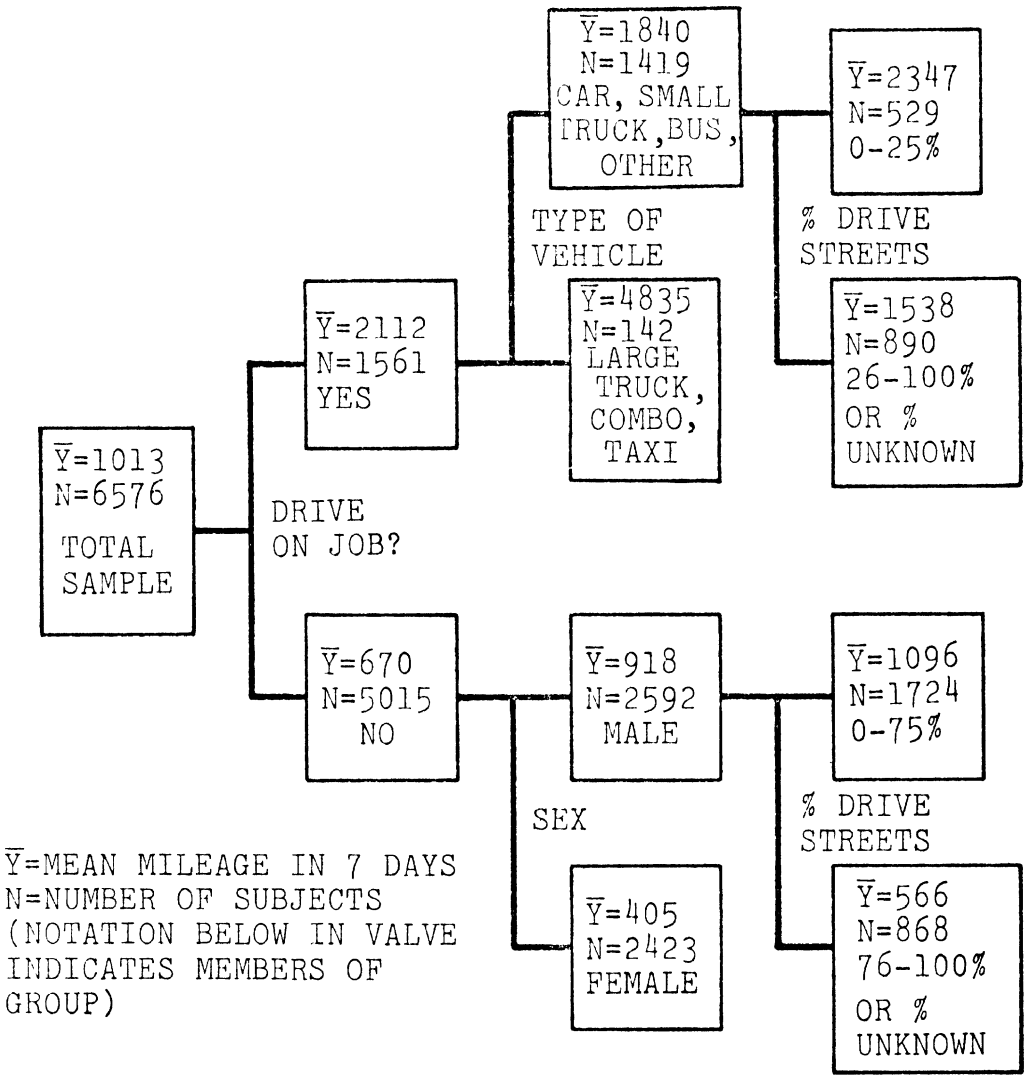

FIGURE 1 Basic AID Chart from Pilot Survey

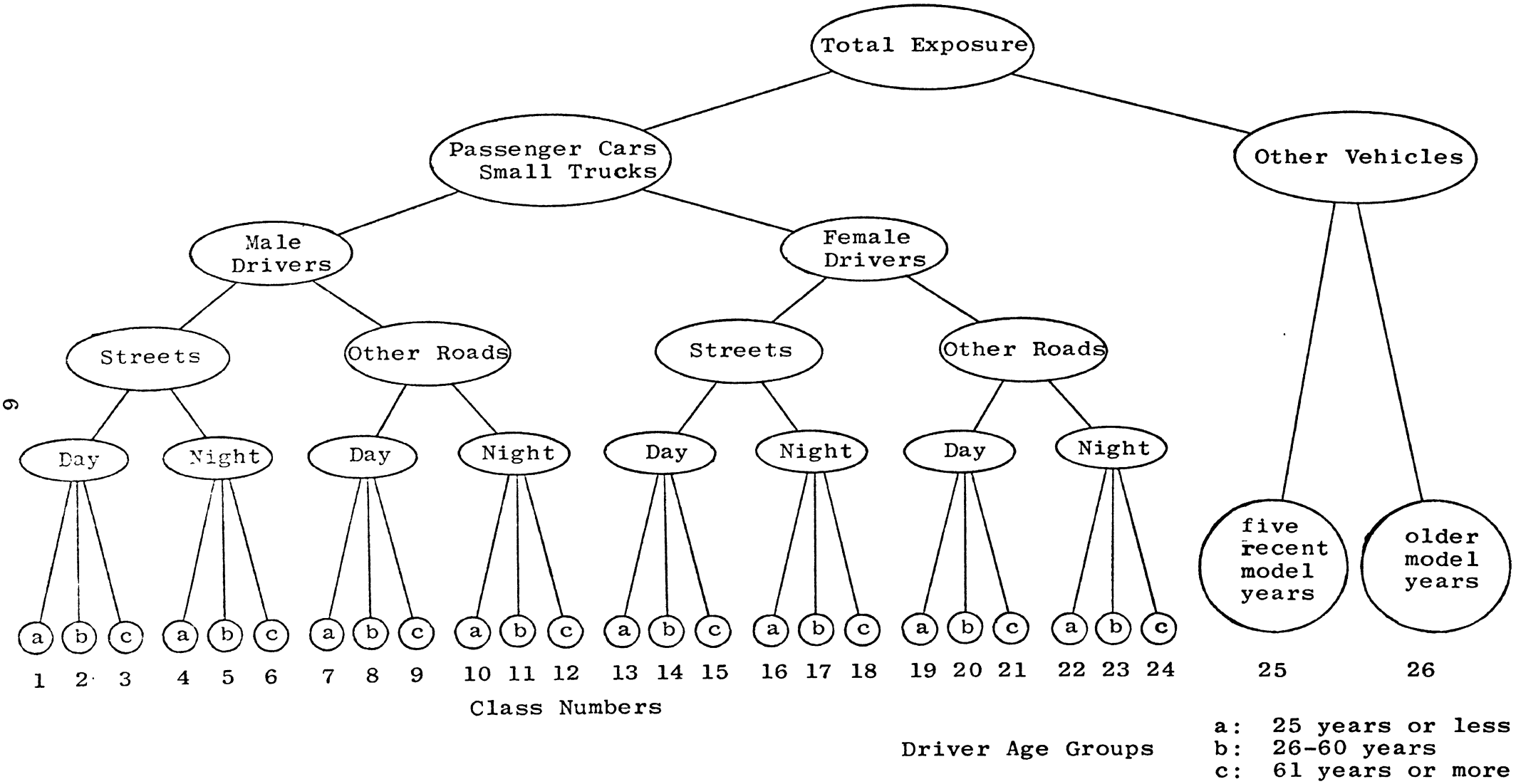

FIGURE 2

Chart of Unique Exposure Classes

Driver Age (3 groups: up to 25, 26-60, over 60)
Vehicle Model Year (five recent years vs. older years)

It is recommended that the six variables be included in future exposure surveys, and that the 26 classes be used to determine total exposure and accident rates.

EXPOSURE SURVEY PROCEDURES

The requirement of this task is to "determine and analyze procedures for exposure sample surveys to provide estimates of vehicle-mileage" for the driver-vehicle-road-environment classifications determined in Task 1.

The approach was to systematically derive a set of feasible alternatives, i.e., basic procedural plans for potential exposure surveys, and to compare the alternatives for accuracy and response rate by means of brief test surveys.

Initial analysis of federal highway safety research needs concluded that operational exposure surveys should be officially sponsored, national in scope, and annual in frequency. Analysis of a scenario of exposure as a travel process among elements of the highway transportation system led to a conclusion that drivers are the only feasible source of accurate, well-classified exposure estimates. Further, the potential magnitudes of driver surveys made it clear that small samples of drivers must be used.

Potential methods of data collection from drivers were structured as follows:

Mode of Collection	Mileage Estimation Method
1. Office Interview	a. Estimate of gross mileage
2. Office Questionnaire	b. Estimates of mileages in several categories of driving or components
3. Mail Questionnaire	c. Estimates of trip mileages after reconstruction of recent trips
4. Telephone Interview	d. Recording of mileages in a trip log
5. Home Interview	e. Recording of cumulative odometer readings

Of the 25 possible mode/estimation combinations, 14 were compared by small specialized surveys with a total sample of about 900. Of the remaining 11 combinations, six were not feasible and the others--home interviews--were rejected because of high cost.

In one survey, odometer readings of the sample were compared with their own gross estimates in an office interview. Odometer reading responses by mail were about 50%, and mean mileage results were in error by only 4 percent.

In the other surveys, independent driver samples were taken in the same county. Exposure rates of about 90% were achieved for all office interview methods and office questionnaire methods, except those requiring a mailed trip log following the interview. Most of the mailed questionnaires and mailed trip logs had response rates of about 50% or less. The telephone response rate was 58%. In most cases, the rates could be raised to an acceptable 80% level by means of one or two follow-up contacts.

Accuracy comparisons were made with the mean value of trip log responses as a reference (because of intrinsic accuracy in one-trip odometer readings). Mean values were within 15% by all methods except the total of component estimates, which reflected the compounding of overestimation.

Although the results are not statistically significant, they do show that gross estimates are probably closest to the more accurate trip log method.

SURVEY COSTS AND RECOMMENDED PROCEDURES

The requirements of this task are to determine comparative costs of alternative survey methods, and then, in conjunction with accuracy findings of Task 2, to "recommend procedures that will best fulfill Bureau requirements."

Cost estimates were then made for 19 of the alternative survey methods, broken down into about 20 cost elements under categories of planning, preparation, data collection, and analysis.

Based on a sample of 5000, cost per data case ranged from about $4.50 to $11.00. Mailed questionnaires were cheapest (up to $4.70); office interviews, office questionnaires and telephone interviews were in the $6 - $7 range; and home interviews were most expensive (nearly $11.00).

At the present time, the office interview and office questionnaire methods are not feasible because many of the states do not require drivers to appear in person at a licensing office when they apply for renewal. Home interviews and telephone interviews are eliminated because of high costs and sampling difficulties. Thus, the mailed-questionnaire type survey is recommended, with attached trip logs as the mileage estimation method.

FIELD TEST PROCEDURES

The requirement in this task is to "recommend field tests to evaluate procedures developed" in the preceding tasks.

A field test of an exposure-survey plan would require real-world implementation of all aspects of the plan, but it would be limited in time and geographic scope. Its purpose would be to validate cost and accuracy estimates of the plan, to discover operational problems, and to evaluate overall performance.

An outline of field-test procedures for a national, mailed-questionnaire exposure survey was generated, including scheduling, questionnaire development, liaison, sample design, clerical preparation, data handling, and analysis. In each part, subtasks for auxiliary procedure evaluations were included.

Procedure evaluations in all parts of the field-test plan appear to be straight-forward, and capable of solid verification of the operational survey plan.

INDIRECT EXPOSURE MEASURES

The requirement in this task is to develop indirect measures of exposure (i.e., substitutes for vehicle-mile data) in situations where it is impossible or uneconomical to obtain vehicle-mile data.

The following indirect exposure measures were analyzed: gasoline sales, one-time-only odometer readings, population, vehicle registrations, roadway right-of-way mileage, and auto insurance premiums. All of these measures are strongly related to vehicle miles, in aggregate. However, they are all incapable of being broken down into classifications according to the six recommended variables of Task 1.

Other problems include gasoline losses, imprecise miles per gallon data, gasoline transfer between states and year to year, vehicle age biases in samples of odometer readings, and time lags and other biases in insurance premiums.

Costs of obtaining indirect exposure data in recommended classifications would not be less than costs of direct exposure surveys; accuracies would not be better. Thus, indirect measures do not provide cost-effective alternatives to direct exposure measures.

RECOMMENDED EXPOSURE SURVEY PROGRAMS

The requirement of this task is to synthesize the findings of the preceding tasks -- i.e., the needs for exposure surveys and the efforts required -- and to make "recommendations for future exposure data collection programs."

The basic findings were as follows:

1. Comprehensive exposure data is needed in highway safety research to permit calculation of accident rates as the key measure of effectiveness.
2. On the basis of need, official exposure surveys should be conducted.
3. Future exposure surveys should use estimates of vehicle miles of travel as the measure of exposure.
4. Independent variables should include vehicle type, driver sex, road type, light condition (day, night), driver age, and vehicle model year.
5. The six independent variables should be used to define 26 unique classifications of exposure, i.e., driver-vehicle-road-environment combinations.

6. Future exposure surveys should be national in scope, on an annual basis.
7. Drivers should be the source of exposure estimates.
8. Small random samples of drivers are adequate for exposure surveys, and necessary economically.
9. The basic mode of exposure data collection should be by means of mailed questionnaires, which have the lowest relative cost.
10. The basic method of drivers' vehicle-mile estimation should be by means of trip logs of one-day duration, which have the highest relative accuracy.
11. Field tests of a recommended exposure survey plan are feasible and desirable prior to full-scale operational implementation.
12. There are no available indirect measures of exposure which are preferrable, on a cost-effective basis, to direct measure of vehicle miles.

In the process of synthesizing these findings, the following conclusions were reached:

1. Eventually, national exposure survey programs should be conducted on a state-by-state basis, so that each state may apply measures of effectiveness to its own unique set of highway safety countermeasures.
2. Official sponsorship authority of future exposure survey programs should be held by the National Highway Safety Bureau.
3. Implementation responsibility of future exposure survey programs should be held by the National Highway Safety Bureau.

Five future exposure programs are recommended:

1. Field Test Program - A nationwide mail survey of driving exposure in the calendar year 1972. Quarterly mailings would be sent to randomly selected drivers in all states, distributed by random selection of each of the seven types of day of the week. State subsamples would be proportionsl to driving population. Total sample size would be limited for economy to about $5,000 such that statistically significant results would be obtained in the 26 unique exposure classes, nationally but not for each state.

   Auxiliary survey methods could be tested simultaneously to provide data on possible evolution of survey method to home interview or office interview procedure.

A first year mail survey program would cost about $250,000, and additional testing of alternatives could cost up to $125,000 more.

2. Operational Exposure Survey Program - Annual surveys, starting in 1973, including modifications derived from the field test program. State subsample sizes would be increased to provide significant results for the unclassified aggregates in each state and for many of the 26 unique classes within each state. Survey designs would continue to evolve as new insights were obtained from yearly re-evaluations.

   Costs of an operational exposure survey would be about the same as the first-year field test ($250,000 annually) if sample size remained at 25,000. A maximum annual cost of about $500,000 is estimated for eventual samples of 100,000 (0.1% of all drivers).

3. Continuing Survey-Evaluation Program - Data from the field test and operational programs should be analyzed continually to determine new variables and exposure classes. This program would involve continuing research, in conjunction with analysis phases of the operational programs, at modest additional cost.

4. Auxiliary Indirect-Exposure Program - Although the potential in the indirect-exposure area is not highly promising, it is likely that gasoline sales data and odometer data will continue to be collected. Also, induced exposure data, derived solely from improved accident data, is still considered worthy of further investigation. Therefore, it is recommended that indirect-exposure research programs be pursued independently at appropriate times, and that the results be compared with results of operational exposure surveys.

5. Other Exposure Sources - Direct and semi-direct exposure data may be obtained opportunistically by means of driver estimates or odometer readings at the time of licensing, vehicle registration and inspection, and accident reporting. Though these exposure sources may not be capable of driver-vehicle cross-classification they may serve as partial checks on direct vehicle-mile surveys. It is recommended that they be considered for inclusion in future revisions of highway safety program standards.

SECTION 3
ACCIDENT DATA INACCURACIES

This section summarizes the results of Phase II (Information on Non-Fatal Crashes), as detailed in Volume II.

ANALYSIS OF ACCIDENT DATA INACCURACIES

The requirement of this task is to "analyze current sources of crash and injury statistics" in terms of their reliability and usefulness in estimating true frequencies of highway accidents and injuries. Emphasis was placed on the study of accident data biases due to underreporting of certain kinds of accidents. Interviews were made of accident experience of drivers, and results were compared with official accident records of the same drivers. Also, injury records were compared with hospital diagnoses of accident injuries to indicate internal inaccuracies of accident reports.

There is a considerable variation in accident reporting requirements among the states. Only 36 states require immediate notification of police in case of a traffic accident. However, all states require "financial responsibility" reporting in case of a traffic fatality or injury, and most require it when damage exceeds a statutory amount (up to $250). In practice, all states provide a form for police reports, regardless of statutes, and all have policies--if only unwritten--to investigate as many traffic accidents as possible within manpower constraints. However, because of these constraints, a great many accidents go unreported, and there is considerable non-uniformity among states as to the degree of underreporting within states. Finally, there is a great deal of non-uniformity of the accuracy of reporting by police agencies both among the states and within them. In mass accident data which is collected, statistics are in error due to both random observational error and systematic underreporting biases.

In comparison of interview accident data versus drivers' official records, frequency of accident was determined in three categories: those accounted for in both survey and records, those in survey but not in records, and those in records but not in survey. There was no way to determine actual accidents not in survey or records. Approximately 35% of all survey accidents were in driver records, and 33% of drivers' recorded accidents were recalled in the survey. For injury accidents, tow-away accidents and high-damage accidents, the percentage of reported accidents were much higher.

In the analysis of accuracy of injury reporting, police injury codes were compared with an AMA injury scale derived from hospital records in Washtenaw County. The most significant finding was the low incidence of severe injuries (AMA scale) in cases which were indicated as severe in the police codes. Only 15.5% of the victims with the highest police code had a really severe injury according to the AMA scale. In addition to these inaccuracies, a sample of police injury codes in 17 states revealed a glaring inconsistency in the distribution of injury codes among states.

CORRECTIONS FOR ACCIDENT DATA INACCURACIES

The purpose of this task was to determine "methods for elimination of effects of major biases and inaccuracies in current information." General recommendations include standardization of accident reporting criteria, simpler report forms, using auxilary data (e.g. licensing information), improved data processing and improved police training for accident reporting. A potential quantitative method of eliminating underreporting bias is the method of the preceding task, i.e., determining ratios of underreporting through sample surveys of drivers' accident experience and comparisons with official records. The derived ratios could be used to extrapolate corrected accident totals from counts in

the official records. However, the survey and record searching costs would be quite high with respect to the normal costs of accident data analysis, even on a sampling basis. This would be especially true if the procedures included derivation of separate ratios for each of the exposure classifications recommended in Volume I or if performed in each state. Further study of potential simplifications of these procedures is recommended.

HOSPITAL RECORDS OF ACCIDENT INJURIES

The purpose of this task is to "determine the feasibility of using hospital records for estimating number and severity of serious injuries." The source of hospital records considered most applicable to this task is the Commission on Professional and Hospital Activities (CPHA), which summarizes about 30% of all U.S. hospitalizations annually. Their past data has included indicators of highway accident victims. However, it does not include emergency treatment in which victims are not subsequently admittted as in-patients. CPHA data on the length-of-stay in a hospital appears to be the best measure of accident severity. This data could be correlated with samples of police injury codes in order to derive corrections in both frequency and severity of injury. In future standardized accident reports, injury codes will be improved. At that time, consideration should be given to detailed procedures for the use of CPHA records and samples of emergency-room records in correcting injury data.

SECTION 4

PROCEDURES FOR AN EXPOSURE SURVEY

This section summarizes the results of Phase III (Driving Exposure Survey), as detailed in Volume III.

DETAILED EXPOSURE SURVEY PLAN

The requirement of this task is to "prepare a detailed sampling plan and procedures for a nationwide driving exposure survey," including schedule, resources, and costs required. The approach was to begin with recommendations in Volume I for a field-test exposure survey using mailed questionnaires with trip-record forms as a base design. Then, the design details were expanded in a straightforward manner to an extent that would be sufficient for a competent survey organization that does not necessarily have expertise in the technical aspects of exposure.

Seven categories of procedures were identified: liaison, sample and survey design, scheduling, questionnaire, data management, data collection, and data reduction. These categories cover all of the work that must be done from the beginning of the survey program to the point that data is ready for analysis.

Liaison includes work with the Bureau of the Budget to secure their approval of the questionnaire and liaison with officials in each state to obtain access to driver lists for subsequent sampling. Sample construction includes distribution of the 28,000 sample among states according to driver population, randomly selecting the proper number of names in each state, randomly selecting the 28 sampling dates (one of each type of day of the week in each quarter of the year), matching name to dates (1000 each date), assigning identification numbers, and obtaining addresses. Scheduling considerations include lead time for planning and questionnaire approval and materials preparation, 13 months for data

collection (final month for followups on December mailings), and a period for data analysis and reporting. Questionnaire development includes design of format and wording to elicit maximum response, writing of cover letters (initial and each of two reminders to be spaced two weeks apart), printing and storage. Data management involves log books, filing system and progress charts to keep track of all questionnaires. Data collection includes assembly and mailing of questionnaires, logging of receipts, identification of non-respondents, and reminder mailings. Data reduction includes checking and coding of individual questionnaires, keypunching, and building of a magentic tape file.

The program is 22 months in duration. It requires about 14 man months of effort, office facilities and materials, and a large computer. The approximate cost is estimated at $253,000.

EXPOSURE SURVEY INSTRUCTIONS

The requirement of this task is to "provide documentation so that the organization which actually conducts the survey will have all the instructions needed to perform the job." The results are simply a detailed outline of direct instructions, within the seven categories above. About 100 specific instructions are derived. Many of the instructions are in sequences that must be repeated for each of the 51 sampling areas. Examples of questionnaires, letters and charts are included.

ALTERNATIVE EXPOSURE SURVEY METHODS

The requirement of this task is to "determine means to check the resultant exposure estimates by alternative collection methods." Several alternative methods were reviewed, based on tradeoff data in Volume I. A national home interview method is recommended as an alternative plan, where exposure data is obtained as a gross estimate, i.e., a single value of mileage for a past time period, e.g., three months. This method is selected primarily because of

the potential low costs if it is attached to an existing national survey. The other alternatives considered (independent home interviews, license office interviews, independent mail surveys) would have to be limited in scope because of sampling costs, and they could be compared only with certain state subsamples from the primary survey.

EXPOSURE DATA ANALYSIS PLAN

The requirement of this task is to "prepare a detailed plan for analysis of the data to be collected," including resources and costs. The analysis plan would be an integral part of the total survey program, and would follow directly from the end of the plans prepared in tasks 1 and 2, i.e., completion of the magnetic tape data file.

From the 28,000 drivers in the sample, an 80% response would yield about 22,400 data cases, sorted in the data file according to the 28 survey dates. Each data case includes a number of trip-mileage estimates, with six independent variable levels as classifiers. The purpose of the data analysis is to computer, from the data file, estimates of mean and total national mileages, in aggregate and in 26 driver-vehicle-road-environment classes.

The initial part of the analysis plan is to prepare a computer program to produce the required exposure statistics. The minimum required computer operations are derived. Additional computer programs may be derived for further classifications of the data (e.g., by quarter of year or by state), for re-evaluation of the 26 classes, or for accuracy checking by an alternative data source. In addition to writing, checking and testing the computer programs, tases of the data analysis plan include performance of computer runs, comparisons and interpretations of results, and final tabulations.

The basic data analysis will be performed during the three months following the end of data collection, and will require approximately 10 man months of effort.