# APPENDIX I

## A DISCUSSION OF DATA GATHERING SYSTEMS

Edwin A. Kidd
CALSPAN CORPORATION

January 16 & 17, 1975

## INTRODUCTION

A data sampling plan that provides an accident data file that is representative of the national population is important. A system for data processing, storage and retrieval to allow early determination of trends in accident, injury and fatality frequencies is essential. But the most critical problems are those concerned with the collection of consistent, coherent data on individual accident sequences with a volume far exceeding that now available.

For too long, those concerned with accident studies of the effects of safety standards already in force have had to make do with either too small samples of reasonably good data or relatively large samples of data whose content is inadequate for the purpose. In the first category is the data bank (and "bank" is too grandiose a term) that has resulted from the individual federal teams of multidisciplinary, professional investigators. These teams can serve useful purposes in special studies, in discovery of problems that would otherwise go undetected and, particularly, in the area of accident causation. By their very nature, they cannot provide a sufficiently large data sample relevant to the implementation of standards aimed at injury and fatality reduction without excessive expenditure of funds.

In the second category are the presently available state data banks of relatively low content data obtained through the use of routine police and driver reporting. These data have been valuable in demographic studies, in the broad-look definition of trends and in statements concerning the magnitude of the overall problem, primarily in fatality frequency. In most cases, such data is totally inadequate in content and precision and, despite the relatively large numbers available at relatively low cost, cannot adequately define injury and fatality reduction resulting from standards implementation.

There is a third category of data collection systems that has evolved over the past few years that lies between the very detailed team approach and the routine police reporting as established independently by the states. The potential exists with this multi-level approach at selected centers around the country (present examples include Calspan in New York, HSRI in Michigan, and HSRC in North Carolina) for a combined data bank that would be a major step toward the attainment of a greatly increased sample size with, and most important, accurate individual accident data with the content required for the purpose.

proper utilization of the potential of these data centers can be realized only if investigator and accident reconstruction aids are implemented that will allow the police to obtain the necessary information with orders of magnitude improvement in accuracy. Local and state police already have the charter to investigate accidents. There are no unsurmountable problems in providing them with the new tools that have been developed for collecting the data that would be the basis for a national data bank sufficient for NHTSA needs in surveillance and effectiveness studies.

## Data Requirements

The list of specific data elements in each accident that are deemed to be essential can hardly ever be complete for the serious analyst. However, the routine and continuous collection of accident data can be tedious, time consuming and costly. Every effort must be made to keep the data requirements to a sufficient set commensurate with the need.

Such sets have been defined a number of times for various ongoing studies. The one presented in Figure 1 is stated in somewhat general terms as it is required, in this instance, primarily for the comparison of data gathering techniques.

| DESCRIPTION OF | INDICATION OF DETAIL |
| --- | --- |
| GENERAL ACCIDENT INFORMATION | SINGLE OR MULTIVEHICLE, RURAL OR URBAN, HIGHWAY CATEGORY, PROPERTY DAMAGE ONLY OR INJURY, OBJECT STRUCK, OVERALL SCENE DESCRIPTION, ROAD SURFACE, AMBIENT CONDITIONS |
| SPECIFIC VEHICLE INFORMATION | MAKE, MODEL, VIN, IMPACT DIRECTION AND DEFORMATION (VDI OR IMPROVED EQUIVALENT), AVAILABLE RESTRAINT SYSTEM, LOADED WEIGHT, TIRES, INOPERATIVE SYSTEMS PRIOR AND AFTER IMPACT |
| OCCUPANT INFORMATION | SEX, AGE, HEIGHT AND WEIGHT, INJURY (MEDICAL REPORT), SEATED LOCATION, USE OF RESTRAINT SYSTEM |
| DRIVER INFORMATION | DETAILS AS IN OCCUPANT ABOVE PLUS DRIVING EXPERIENCE, TRAINING, CONVICTIONS, PHYSIOLOGICAL CONDITION, PSYCHOLOGICAL INDICATIONS, ACTIONS PRIOR TO AND DURING ACCIDENT SEQUENCE |
| IMPACT ENVIRONMENT | SPEED AT IMPACT, RESULTANT SPEED CHANGE TIME HISTORY, COMPARTMENT DECELERATION |

**Figure 1 ACCIDENT DATA REQUIREMENTS**

Even this rather simplified listing appears formidable. However, to some degree each of these elements or approximations thereof are being obtained, by one means or another, by some of the present ongoing programs. There is no element in the outline presented in Figure 1 that is not germane to existing standards. If we settle for a system that provides accurate information less than this and/or only for a quantity of a few hundred or even a few thousand cases, NHTSA cannot do the job it has been directed to do.

Definition of the total number of accident cases required annually for an adequate national data bank can be made if (1) the questions to be asked of the system can be identified both for the present and future; (2) the accuracy with which the particular data elements can be measured is known or can be appropriately approximated; and (3) the statistical analysis techniques to be employed can be agreed upon. This is not meant to imply that such analyses and decisions should not be made. However, there are no statistical procedures that can adequately overcome the past and current inaccuracies with which such extremely important data elements as impact speed and speed change have been reported if they have been reported at all.

There will need to be a parallel effort of statistical analyses to indicate what questions can be addressed with acceptable statistical significance as a function of particular sample sizes along with the determination of the funds that can be made available. The financial impact of standards on the consumer has been and will be considerable — billions of dollars annually. Figure 2 presents an average cost per car for the FMVSS to date based upon idividual autombile manufacturer’s data. It seems prudent to schedule funding for the primary surveillance effort -- accident data collection -- commensurate with the far reaching decisions that depend upon such data.

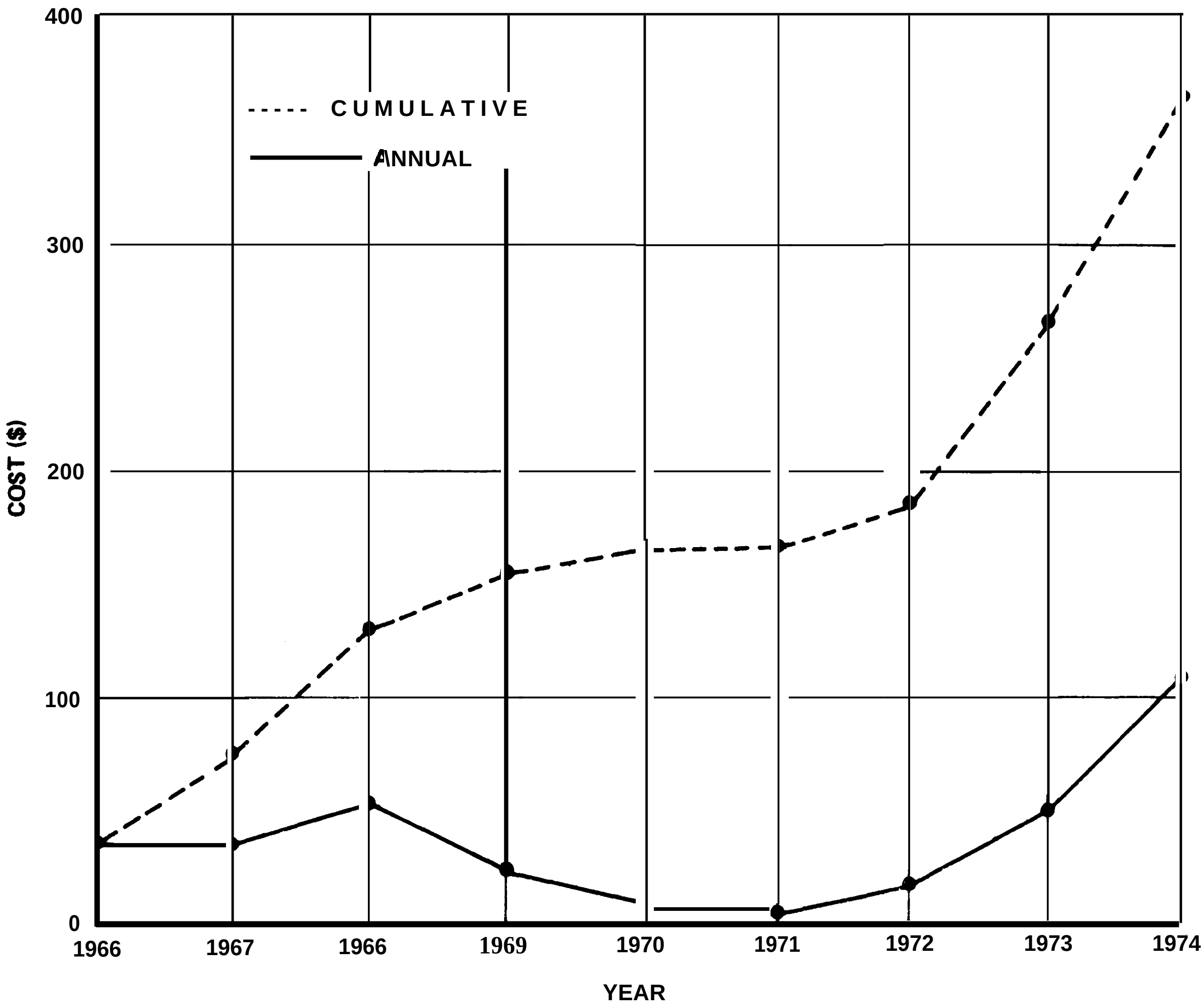



Figure 2 AVERAGE COST IMPACT OF FMVSS ON AUTOMOBILE PRICE

## Data Gathering Techniques

In order to obtain the data required on each accident, every accident analyst would gladly utilize whatever data gathering techniques are available. Ideally, crash recorder information, police and driver reports, intensive investigation team reports and on-scene reconstructions of the accident through computer aids to investigators would all be gratefully accepted by every serious analyst for each accident. In fact, no analyst would refuse any available high speed photographic coverage (in color, of course) .

Obviously, it is neither practical nor essential that all of these systems be provided for the achievement of the basic national data bank. It has already been stated that the intensive investigation teams may have other purposes but cannot provide the data in the quantity required. It has also been noted that existing state data, comprised of merged police report, vehicle registration, and driver licensing files do not provide the content required for the evaluation of safety standards.

## Crash Recorders

Crash recorders can only provide a portion of the desired information as a supplement to continuous accident investigations. At best, a recorder can provide only the information outlined in Figure 1 under "Impact Environment" plus driver control actions during the accident sequence and an identification of the vehicle in which it is installed. Despite the fact that the information a crash recorder is designed to obtain is the impact environment, and that this is the data now totally lacking or sadly inaccurate, a detailed accident investigation would still be required to provide the essential general accident, specific vehicle (including the other vehicle) and occupant and driver information. Thus, the overall cost of an accident investigation would include essentially the present costs plus those associated with the provision of crash recorders.

The numbers game must be considered as well in the consideration of crash recorder Installations. The actual number of accidents that would be available for analysis would be a marked attenuation from the total number of crash recorder installations (Figure 3). Assume that recorders were installed in 100,000 automobiles. No more than 1 in 4 of these automobiles would be involved in any sort of accident annually. This reduces the number of accident cases with crash recorder information to no more than 25,000. If it is further assumed that the accidents of principal interest are those of more than minor severity, for example, tow-away accidents (approximately 12.5% of all accidents), then only 3125 accidents would be available. If we examined only the highest volume model of the major American manufacturer (approximately 25%) the number of accidents available would be approximately 781. Further division of these accidents into accident type, direction of impact, etc., would further diminish the numbers. This severe attenuation would be greatly increased for car make and models other than the one with the greatest penetration of the market.

It is recognized that the crash recorder is designed to provide crucially important information on impact environment that has not been otherwise available, at least in quantities with acceptable accuracy, However, there is now available another method, as discussed below, for obtaining this information with accuracies that appear excellent. Both methods should be compared in staged crash tests and considered for some possible joint use as mutually reinforcing data sources. However, the computer aided system, with its outputs of a detailed scene description and an accurate reconstruction of the accident, offers the most promise, as a fundamental element of a continuing data gathering system.

| | |
|---|---|
| ANNUAL CRASH RECORDER INSTALLATIONS | 100% |
| VEHICLES INVOLVED IN ANY ACCIDENT | 25% |
| VEHICLES INVOLVED IN SEVERITIES > TOW AWAY | 3.125% |
| MAJOR MANUFACTURERS HIGHEST VOLUME MODEL | .078% |

Figure 3 CRASH RECORDER EQUIPPED AUTOMOBILES IN ACCIDENTS

Computer Aided Investigation and Reconstruction

With support from NHTSA, **this** on-scene accident investigation and reconstruction system has been developed and demonstrated (Figure 4), An automated range-finder transit with associated computer hardware and readout (Figure 5) provides a drawing of the accident scene and supplemental accident information as required (Figure 6). These physical evidence data are transmitted via a radio-telephone link to a centrally located computer which returns a reconstruction of the accident (Figure 7).

In actual reconstructions of staged accidents, this investigator's tool has faithfully reproduced the accident sequence with impact speed and speed change reconstructions of 2-3% accuracy. With this system, police investigation teams can generate high quality accident data in the course of performing their normal police functions. Yet the system has been found, during field trials by police personnel, to actually ease the tasks of scene measurement and reporting. Thus, both the users of accident data and the police can benefit from adoption of this system.

The economics of adopting the system would be extremely attractive from the viewpoint of elimination of labor costs in the generation and reporting of accident data for research purposes. The end product is already in digital format for statistical analysis.

The nature of the output from the van also lends itself directly to a central data bank or regional data banks receiving reconstructed accidents and supplementary data over existing telephone lines. This continuous updating of current data is particularly attractive. At present, the best a state can do, those few that can supply merged accident tapes, is provide a year's data six-eight months after year's end. A dedicated data collection center, such as presently sponsored by NHTSA, can provide computer tape updates of collected, augmented police reported data every three months with a three month delay.

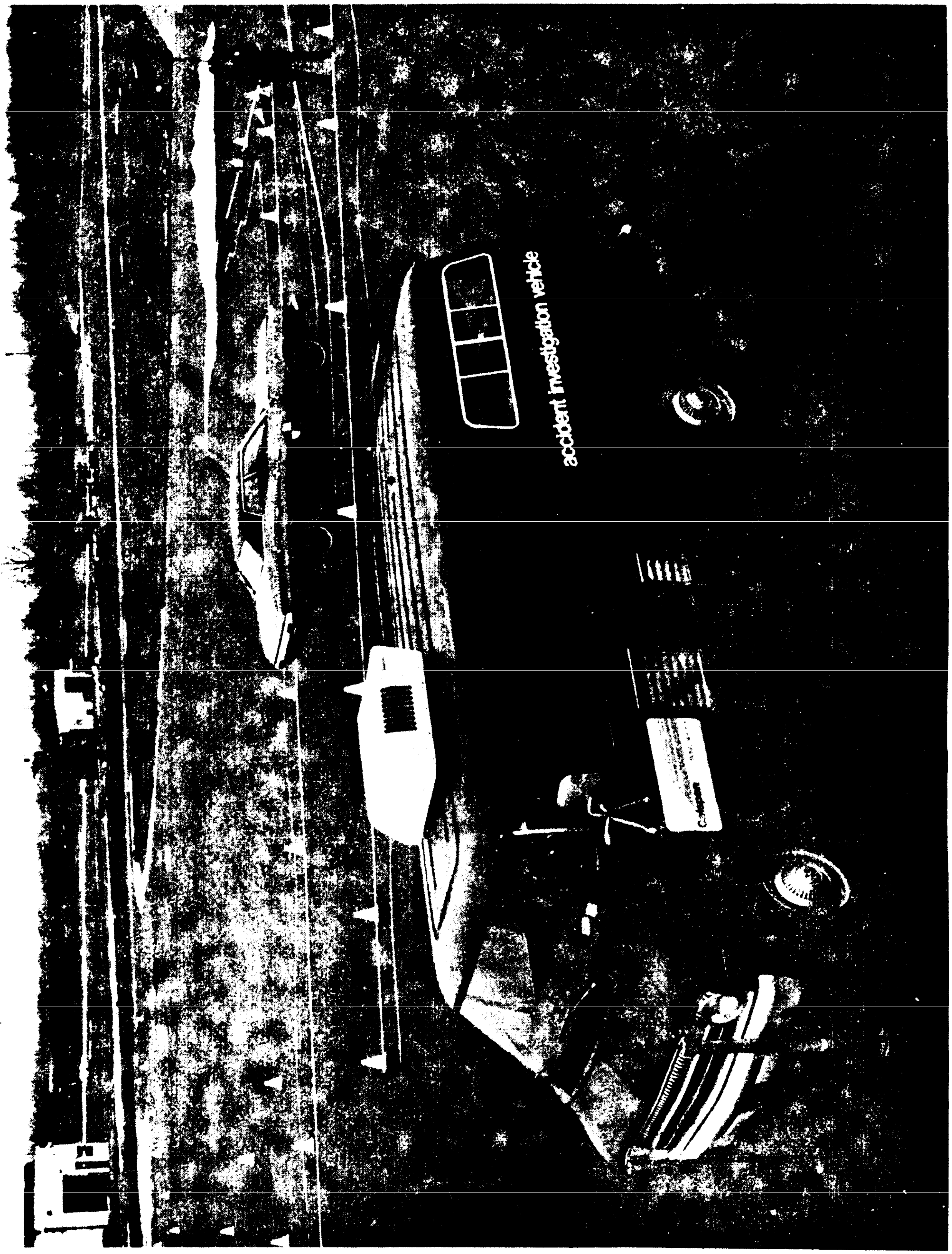



Figure 4 ACCIDENT VAN

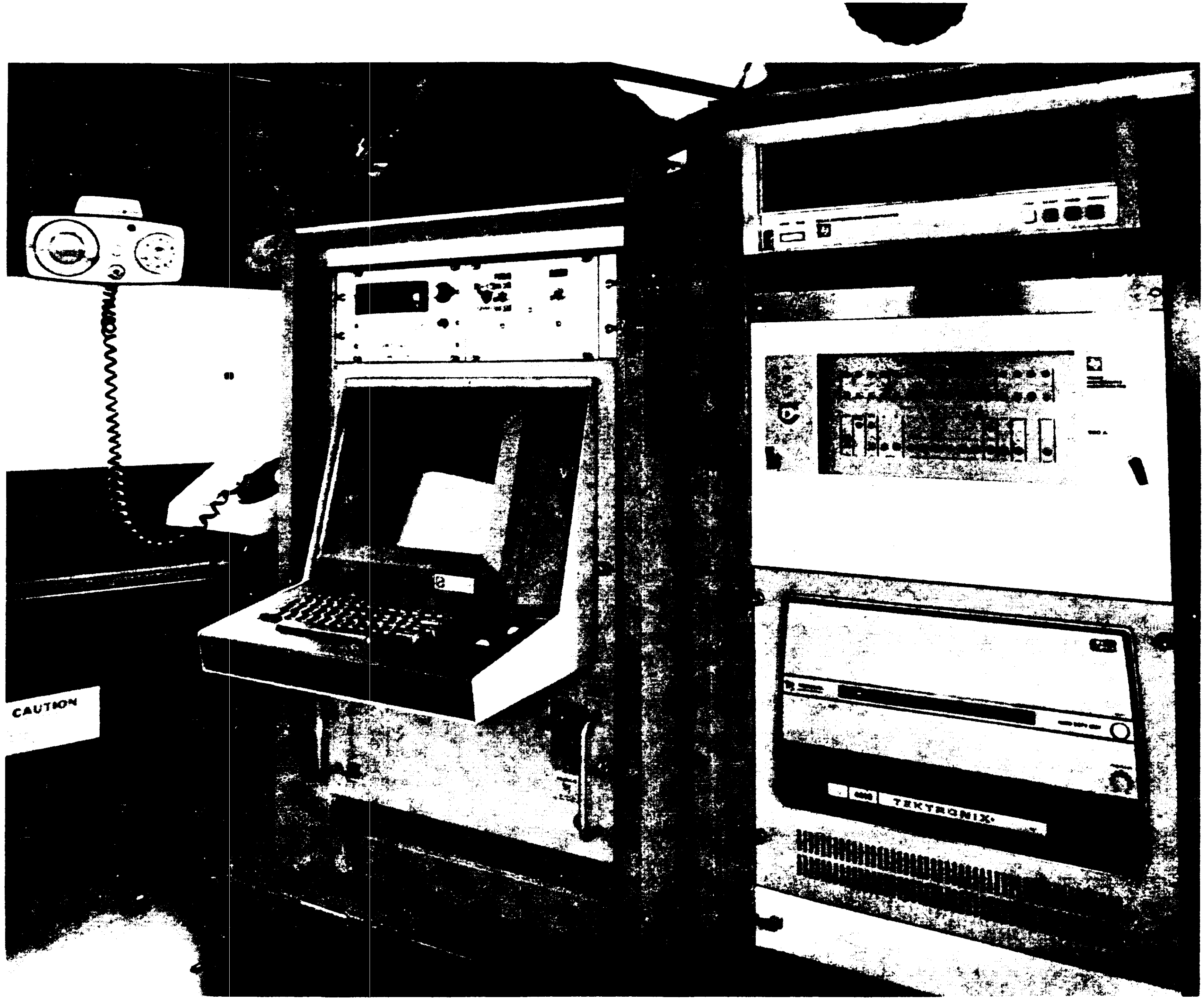



Figure 5 VAN INTERIOR

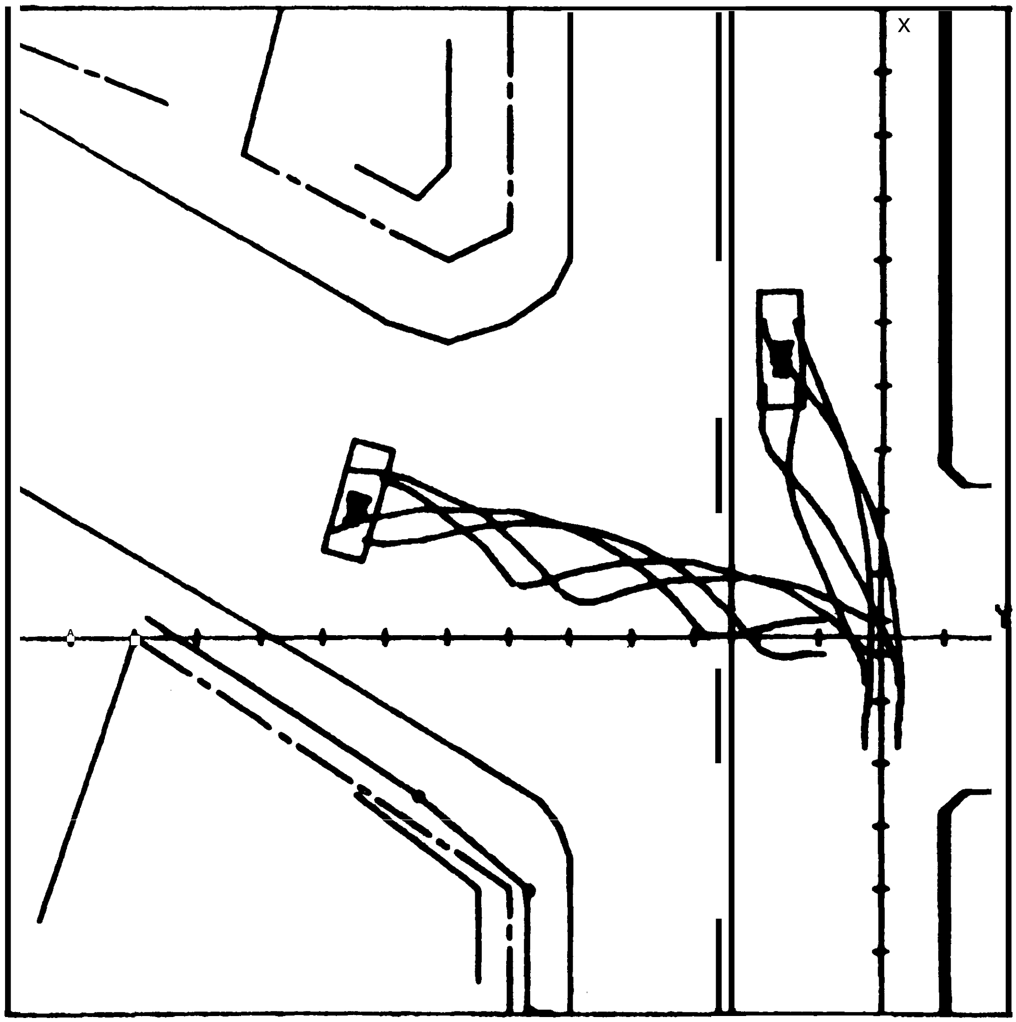
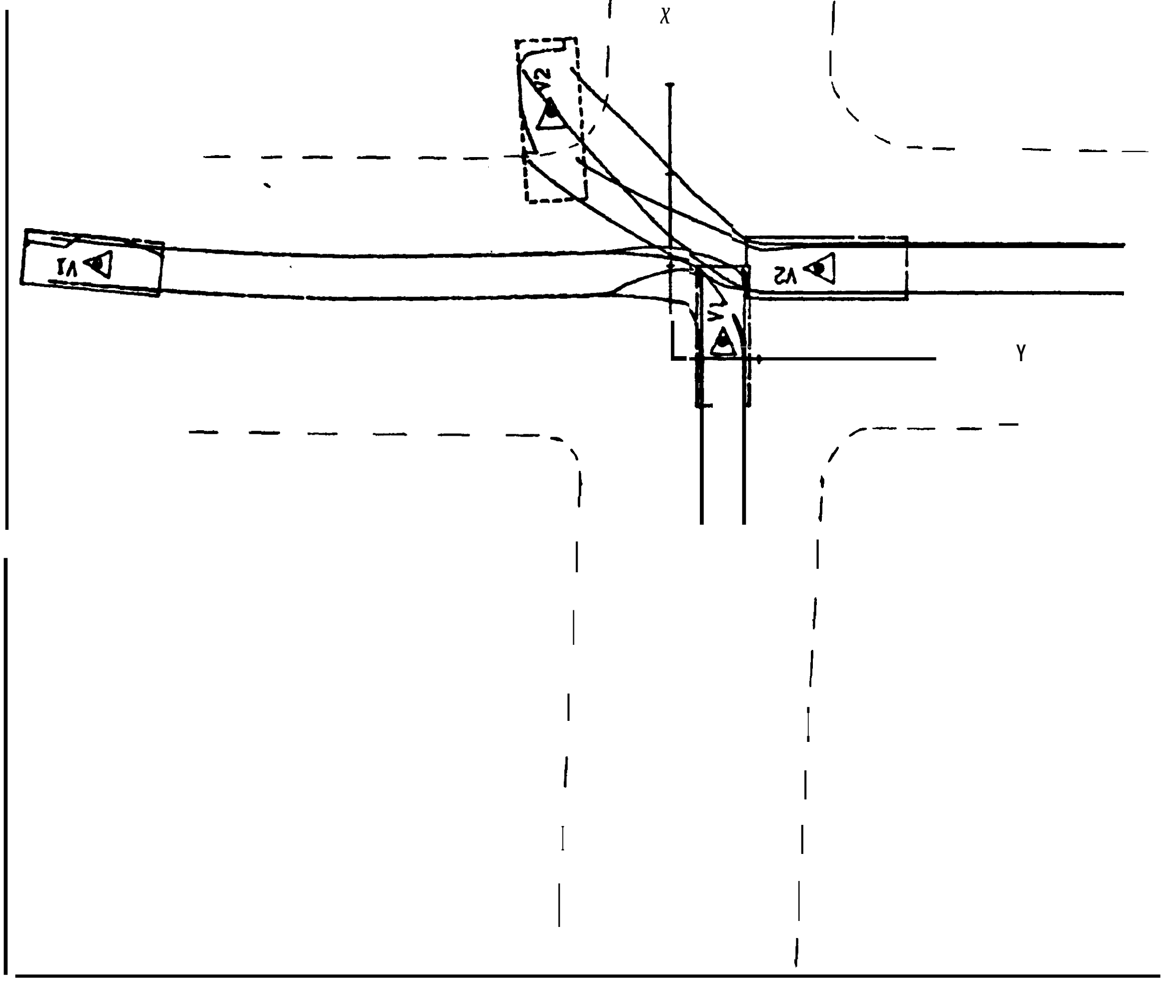



FIGURE 6 SCENE AND ACCIDENT RECONSTRUCTION

GRAPHIC DISPLAY OF OUTPUTS OF-

ACCIDENT RECONSTRUCTION

COLLISION AND TRAJECTORY




AXIS INTERVALS ARE 10. FEET

| RECONSRUCTED POSITIONS AND VELOCITIES AT IMPACT | | | | | | | DISPLAYED FINAL POSITIONS | | | | VEHICLE DAMAGE INDICES | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | C.G. POSITION | | HEADING | | | | C.G. POSITION | | HEADING | REMARKS | | |
| | XC1 | YC1 | PSI1 | FWD | LATERAL | ANGULAR | XC1F | YC1F | PSI1F | | | ΔV |
| | FT. | FT. | DEG. | MPH | MPH | DEG/SEC | FT. | FT. | DEG. | | | MPH |
| VEHICLE # 1 | 1.9 | 6.0 | -0.0 | 24.0 | 0.0 | 0.0 | 10.3 | -63.5 | -84.6 | IN MOTION AT 3.5 SEC AFTER INITIAL CONTACT | 02RYEW4<br>04RZEW3 | 20.8<br>10.2 |
| VEHICLE # 2 | 10.0 | 16.8 | -90.0 | 36.8 | 0.0 | 0.0 | 26.8 | -13.2 | -3.4 | VEHICLE AT REST | 11LFEW2<br>09LPEW3 | 14.9<br>6.9 |

FIGURE 7 SAMPLE ACCIDENT RECONSTRUCTION

## Data Collection System

The mobile accident van will be ready for general police use after some additional field trials. The accident reconstruction computer model (SMAC), as incorporated in the van reconstruction software and hardware, has been distributed and is in extensive use; additional validation for a variety of accident situations is planned. The total system works and works very well. It is appropriate to consider how it might be incorporated in a complete data collection system that would provide greatly improved and accurate accident data with the quantity and content required for surveillance of standard's effects on accident consequences.

strategically located data centers have been suggested for the collection of regional data samples of the multilevel type. An appropriate distribution of accident vans for use by police investigators within each of these regions would provide continuous data into regional data banks and/or to a single data bank.

There are two primary options for van configurations. Police investigators can be equipped with either a Scene Van or a Reconstruction Van. The Scene Van would provide a description of the physical evidence and supplementary accident data (Figure 6). Hard copies of this information in appropriate format would be available as the police report. In addition, all data would go on tape cassettes to be forwarded to the particular data center for reconstruction of the accident by the SMAC model. The reconstructed accident information would supply the center's data bank. Appropriate retrieval and analysis programs would provide immediate analyses as required.

Reconstruction Vans, with either self-contained reconstruction capability (more comprehensive on-board computing equipment) or radio link to a regional time-share computer or computers would by-pass the step of accident reconstructions at the data center. Also, a more desirable feedback to the accident investigator at the scene would be available. (A successfully reconstructed accident is the best check of the completeness and accuracy of the scene data.) Each reconstructed accident data set would be stored (short term) within the van and transmitted, when convenient, to the data center at available terminal locations (already present at police agencies).

A rough approximation of the cost for two assumed data collection systems is given in Figure 8. These are given to provide an approximate range of system costs for collection of 100,000 cases annually. The cost of a radio link reconstruction van system would be somewhat less than the self-contained van with a resulting overall cost close to that of the scene van system. Final selection among these alternatives should consider, in addition to basic costs, operational factors including the advantages of program updating and modifications with either the Scene Van or the radio-link Reconstruction Van and the overall data improvement that would result from the Reconstruction Van.

Regardless of the system selected, costs per case of less than $100 are estimated. This appears to be quite a bargain. The system would provide 100,000 cases per year for whatever investigation criteria is desired, e.g., tow away cases. Costs per case are essentially independent of data sample size. The assumed rate of cases per van per year is conservative, considering that police agencies operate 24 hours per day.

| | ANNUAL COST - $1,000 | |
|---|---|---|
| | SCENE VANS | RECONSTRUCTION VANS |
| 500 VANS | $2,500* | $3,750* |
| 5 RECONSTRUCTION UNITS (1DEDICATED MINI COMPUTER AT EACH OF 5 CENTERS) | 0* | – |
| SERVICE, MAINTENANCE | 500 | 750 |
| ACCIDENT RECONSTRUCTION LABOR (100,000) | 200 | – |
| TOTAL | $3,210 | $4,500 |
| COST PER CASE 100,000 CASES, 200/YEAR/VAN) | $32 | $45 |

*ASSUMES 10 YEAR USEFUL LIFE WITH LINEAR DEPRECIATION

Figure 8 APPROXIMATE, FIRST-O[illegible] COSTS ACC DENT VAN DATA COLLECTION SYSTEM

Obviously missing from the estimates of Figure 8 are labor costs for data acquisition at the scene. It is assumed that police would do the investigations. With police use of vans for their own investigative purposes with improved efficiencies over the present, acceptance of the vans should be readily realized.

Based upon our experience to date, we believe the usefulness of this mobile system to the police themselves can be demonstrated and there is no concern that they cannot properly operate the equipment. Implementation of a total collection system employing this scene data gathering capability will provide NHTSA with the information needed.