Large Truck Accident Rates: Another Viewpoint

W. D. GLAUZ and D. W. HARWOOD

ABSTRACT

There have been many studies that have attempted to ascertain any difference in the safety of various large truck configurations. Special attention has been paid to double trailer combinations (doubles) compared to tractor-semitrailer combinations (singles). Most researchers have found little if any difference in their respective accident rates. The major exception is a large, widely quoted study conducted by BioTechnology, Inc., for the Federal Highway Administration. It concluded, among other things, that doubles have significantly higher accident rates than do singles. The research reported in this paper consisted of a thorough review of all aspects of that study, with particular emphasis on the structure and contents of the three major data bases used. These included data on 2,112 large truck accident involvements, vehicle classification count data taken at 78 sites, and driver and truck data (including size and weight) on more than 32,000 trucks. It was found that the conclusions drawn by BioTechnology, Inc., regarding doubles versus singles, as well as most of the other major conclusions, are not supported by the data bases. Although the accident data base may be useful to other researchers, if used with care, the two data bases needed to estimate exposure are totally inaccurate, and results derived therefrom are probably meaningless.

The passage of the 1982 Surface Transportation Assistance Act (1) (STAA) by the U.S. Congress accomplished many things. Among these was a mandate that the states not prohibit truck combinations involving two trailers from using the Interstate highways or other highways designated by the Secretary of the U.S. Department of Transportation.

Although such truck combinations, often called doubles or double bottoms, were already legal and regularly operating in many states--notably in the midwest and west--they were not legal in many of the southeastern and eastern states. Their legalization by the STAA raised many questions in the minds of concerned officials and citizens, especially questions about the safety of doubles compared with that of the more common tractor-semitrailers (singles).

The literature contains a number of original studies of this issue (2-10) as well as several reviews and summaries (11-13). Most of the studies found little difference in accident involvement rates of doubles and singles. The major exception to this trend is a study by Vallette et al. (10), often referred to as the BioTech study. It concluded that "twin trailer combinations have a significantly higher accident rate than single tractor-trailer combinations." (Vallette et al. use the word "twin" instead of "double"; however, "twin" is often taken to be the more limited configuration of a tractor plus twin van-type trailers. This distinction will be addressed further at a later point in this paper.) Their findings were met with much skepticism and criticism (14). Nevertheless, they were ultimately "accepted" and widely quoted (11-13).

Among those reviewing and criticizing the study were members of the trucking industry. Subsequently, Midwest Research Institute (MRI) was asked to conduct an in-depth, critical review of the work. This paper covers some of the major findings of that review.

THE BIOTECH STUDY

The work done by Vallette et al. (10) was a substantial and ambitious research effort funded by FHWA. It involved collection of massive amounts of data covering the 18-month period from July 1976 to December 1977, reduction and analysis of these data, and preparation of a final report that was ultimately published by FHWA in late 1981.

The concept was to determine the effect of a number of factors on accident rates of large trucks. The list of factors included highway type, truck weight, truck configuration, truck length, cargo type, and driver age and experience. To calculate an accident rate for a specified set of factors (e.g., doubles on urban freeways) requires knowledge of (a) the number of doubles involved in accidents on such facilities during a stated time period; and (b) the vehicle-miles of travel of doubles on those facilities during the same time period. Mathematically, the accident rate is calculated as

$$R = A/E \qquad (1)$$

where R is the accident rate (e.g., accident involvements per million vehicle-miles), A is the number of accident involvements for trucks with the specified characteristics, and E is the comparable "exposure," usually measured as vehicle-miles of travel. To make the representation more explicit, Vallette et al. (10) expressed Equation 1 as

$$R_{ij} = \sum_k (A_{ij})_k / \sum_k (E_{ij})_k \qquad (2)$$

In this expression i is the roadway class of interest (e.g., urban freeways), j is the variable or set of variables of interest (e.g., doubles or doubles weighing 50,000 to 60,000 lb), and k is a specific site or section of highway.

The research plan envisioned using a stratified random sample of sections of highway in a number of states. Three types of data were to be collected for each section to form data bases. The first type consisted of accident data for all truck accidents, and included information from the investigating officers' reports, plus more in-depth follow-up information. These data provide values for the terms $(A_{ij})_k$.

A second type of data was derived from vehicle classification counts (counts of the number of vehicles of specified configuration such as singles and doubles). The counts were to be made quarterly over the 18-month period, with each count covering a 7-day period, day and night. The primary method used to obtain these counts was a photographic sampling process described later. From these counts the vehicle-miles of travel could be calculated. The latter, in turn, can be used as $(E_{ij})_k$ in Equation 2 for questions involving only the truck configuration.

A third type of data base was required to apply Equation 2 to more detailed questions, such as the effect of truck weight or driver age. This data base, referred to as the size and weight data base, was to be obtained from weigh scales and truck stops. This information was intended to provide the appropriate values of $(E_{ij})_k$ needed to answer the more specific questions.

Altogether the study included 2,112 truck accident involvements at 78 highway sections in six states. A truck accident involvement is any truck involved in an accident (an accident involving two trucks yields two truck accident involvements). The second data base was planned to include 7 days x 6 quarters x 78 sites = 3,276 site-days of classification data. The third (size and weight) data base included information on 32,102 trucks observed during the study.

THE MRI REVIEW

The findings reported in this paper resulted from an unusually detailed review of all aspects of the BioTech study. Initially, the report (10) and its various drafts, along with the contractor's letter reports provided by FHWA, were reviewed. Internal FHWA memoranda and a report by the U.S. Department of Transportation Transportation Systems Center (TSC) presenting an independent review of the study (14) were also reviewed. Then, a number of the sites used in the study were visited by the lead author.

A critical activity was the acquisition and analysis of the data bases. The accident reports assembled by BioTech researchers were obtained and studied. Four data tapes containing the summarized data bases were used. More important, a fifth data tape, heretofore unexamined by others, was obtained from FHWA and decoded. It contained the "raw" vehicle classification count data base. From this data base it was possible to deduce the assumptions made by BioTech in deriving the summary classification count data tape that was later used by FHWA and TSC (14) in their review of the study. Unfortunately, the rawest form of the data (films and recording forms) no longer existed so they could not be examined.

Finally, comparative data (accident, exposure, and weight) were obtained from the California Department of Transportation (Caltrans). These data were then analyzed and compared, where feasible, with like data obtained in the BioTech study.

Although the BioTech study dealt with a multitude of truck-related issues, this review was directed mainly to the doubles versus singles issue. However, other issues that had a bearing on, or might help to explain, the authors' findings on doubles and singles were examined.

FINDINGS

The review process identified a number of research areas that were investigated individually. These ranged from the research design and its implementation to the in-depth study of each of the data bases and to the analytical and calculational procedures used. These individual areas and findings relative to them are described in this section. In general, the discussion of each area includes what the study intended, what the report states or implies was done, and the present authors' determinations.

Research Design--Site Selection

Vallette et al. clearly indicate at the outset that the research is not to be interpreted as nationally representative. Indeed, they purposely selected six states in which to collect data, each of which was in some sense unusual in regard to truck operations. The selected states tended to have unusual legal limits, high truck accident histories, or high truck volumes. One state allowed trucks up to 165,000 lb GVW; and other allowed 105-ft-long combinations. Thus, at best, the results obtained may only be applicable to the areas of the country from which the data were obtained.

Of the six states, only three allowed doubles, and the trucks in one of those three states (Michigan) are unique to that state. Thus all comparisons of doubles versus singles used data from just two states--California and Nevada. Indeed, 93 percent of the 189 accident involvements of "normal," five-axle doubles were from a single state--California. Thus the findings regarding doubles versus singles are perhaps valid only for that state. (As noted later, many of the doubles in California are of quite different configuration than are those observed elsewhere in the nation.)

A sophisticated, stratified random sampling scheme was planned. Six roadway types (later reduced to four) formed one level of stratification; truck accident frequency within roadway type formed a second level. However, final site selection procedures apparently did not follow the plan--high accident locations dominated. Further (as discussed subsequently), because the final analysis ignored the stratified design, the results largely reflect traffic safety at a very few sites. The data given in Table 1 illustrate this effect for the 31 sites in California and Nevada.

Classification Data

Site Uniformity

It was the stated intent of Vallette et al. that the sites would have "well-defined points of entrance and egress: to gain some assurance that vehicles (trucks and cars) entering one end of the site would be exiting the other end." It is evident, on examination of the sites, that this intent was not realized.

Examples are shown in Figures 1 and 2; many others could be cited. Figure 1 shows a 4-mi section of I-80 on the southern and eastern sides of Sacramento, California. The data quoted are from Caltrans vehicle classification counts obtained in 1976. Figure 2 shows a highway section in the agricultural Imperial Valley near Indio, where about half the truck traffic is generated within the section. Another example is Site 141, a section of the San Diego Freeway. The site is intersected by several small interchanges and by a major interchange with the Harbor Freeway, which experiences a heavy volume of turning movements. The cross section varies from three to five lanes northbound and from three to six lanes southbound. These examples, and many others,

TABLE 1 Dominant Sites

Stratum	No. of Sites[a]	Dominant Site	Dominance (%)[b]	Comments
Rural freeway	8	Site 114 I-5 (Grapevine)	56	Long, steep grade north of Los Angeles, eight-lanes, 26,500 AADT, 20 to 30 percent trucks, heavy fog common
Urban freeway	7	Site 145 I-80 (San Francisco-Oakland Bay Bridge)	47	Ten-lane toll bridge, 183,000 AADT, 9,500 trucks per day
Rural nonfreeway	11	Site 123 CA-152 (Pacheco Pass)	43	16-mi winding mountain road, partially two-lane, 12,000 AADT, 20 percent trucks, often cloud obscured, posted for high, gusty winds
Urban nonfreeway	5	Site 152 US-101 (just south of San Jose)	70	10-mi section of US-101 not yet up to Interstate standards, four lanes, uncontrolled access, 40,000 to 46,000 AADT, 8 to 9 percent trucks

[a]Sites in California and Nevada from which the doubles versus singles data were obtained.
[b]Percentage of all truck accident involvements for the stratum.

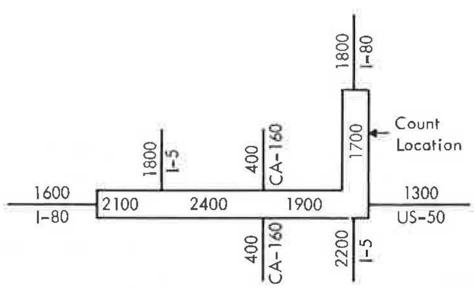

FIGURE 1 Site 143 (I-80) showing 5+ axle truck AADTs and major intersecting highways.

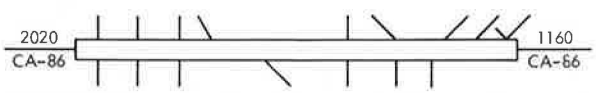

FIGURE 2 Site 121 (CA-86) showing total truck AADTs and intersecting roadways.

show that observations at one point in a site often were not representative of the whole site.

Directional Split

Classification data were obtained by observing traffic in only one direction. This would be acceptable if it could be demonstrated that the volumes, classifications, and accidents in both directions were about the same. However, such is not the case.

Table 2 gives, from the BioTech data, the difference in accident involvements, by direction, for those sites with a reasonable number of accidents during the study period (12 or more) and, for illustrative purposes, which did demonstrate a sizeable directional effect. Note that all four of the dominant sites of Table 1 demonstrate this difference.

Also shown in the table are the directions used for the classification and size and weight data. Note that these data were generally obtained from the direction having the smaller number of accidents.

Vehicle Sampling Adjustments

The classification and, subsequently, the $(E_{ij})_k$ data were derived by adjustments to counts obtained

TABLE 2 Sites with Directional Imbalance in Accident Numbers

	Accident Involvements		Direction Sampled	
Site	Directional Split	Dominant Direction	Classification Data	Size and Weight Data
112	14/8	South	North	Both
114	111/67	North	North	South
115	12/7	South	South	Unknown
123	36/20	West	East	East[a]
141	15/10	South	North	South
144	16/2	West	Both[b]	Both
145	64/39	East	West	West
152	17/13	North	South	South

[a]Both directions were eligible, but predominantly eastbound were interviewed.
[b]Switched directions part way through study.

through a photographic sampling process. A camera was positioned at each site, typically for a week or so at a time. The camera was triggered via an axle detector and counting circuitry designed to take a photograph after every N^{th} axle passage, where N was site specific. It was then intended to adjust these raw sample counts to arrive at estimates of annual vehicle-miles of travel by site and vehicle class.

State average annual daily traffic (AADT) data for the various sites were the basis of the adjustment. The state AADTs were first apportioned into trucks and nontrucks; then the truck values were further divided into numerous truck classes.

The researchers determined that their samples indicated a far higher percentage of trucks than state data showed. As a consequence, it was decided to use the state truck-nontruck split instead of what was directly observed on the films. This correction was made for only 18 of the 31 sites in the two states; inexplicably, no correction was made for the other 12, so their truck volumes remain greatly exaggerated. (At Site 145, the Oakland Bay Bridge, Vallette et al. used manual classification counts instead of the photographic process, so no adjustment of this type was needed.)

The reason for the discrepancy is that the sampling scheme was of axles, not vehicles. As stated by BioTech (10), "a five-axle truck would have five chances of having its picture taken while a three-axle truck would have only three chances of having its picture taken." The report further stated that the raw counts were adjusted to account for this sampling bias.

Although such an adjustment was clearly intended, it was not made. As stated earlier, the BioTech data tape, provided by FHWA and not previously reviewed, was thoroughly analyzed. It contained the raw counts of vehicles in 10 classification categories for each site, observation period (one of six calendar quarters), date, and time of day (day or night). It was possible to mathematically reproduce the published summaries from this tape without making any axle count adjustments. Thus, many-axle vehicles are overrepresented, relative to passenger cars and other vehicles with fewer axles, in the BioTech classification data base.

Amount and Distribution of Data

The research plan called for classification data to be collected day and night at each site for one week during each of the six calendar quarters of the data collection period, and the report indicates this was done (10). However, the raw data tape does not contain 6 quarters x 31 sites = 186 data sets. There are only 74 sets of data from the films, plus one set from Site 145, where manual classification counts were made just once for a portion of one day. No data were obtained in the first quarter, and very little was obtained in the sixth. The bulk of the data (representing 25 of 31 sites) was obtained in the fifth quarter (July-September 1977). About half that amount was obtained in each of the second through fourth quarters. Because no seasonal adjustments were performed, the data base is biased toward the summer and early fall.

There was also a bias in the amount of data by site. Although some data were obtained at each site, no site had more than four data sets; most had just two. There were three sites for which only one data set was obtained, including Site 145.

The data base distinguishes between daytime and nighttime counts. All of the 74 photographic data sets include daytime data, but only 26 contain any nighttime data. Of these, three were of such questionable quality that BioTech discarded them. Review of the data suggested that, relative to the doubles versus singles issue, an additional 16 nighttime data sets could have been discarded either because of their extremely small sample sizes (fewer than five trucks in one of the two categories) or because of poor quality (discussed further subsequently). That would leave only seven sets of possibly useful nighttime photographic data, not the 186 the report implies were used.

When both day and night data existed in a given data set, BioTech combined them in a reasonable fashion to estimate the 24-hr distribution. Unfortunately, when night data did not exist, which was the case in 51 of the 74 data sets available, they simply used the daytime data as representative of the entire period.

This would be acceptable if, for example, the doubles-singles relationship could be shown to be the same both day and night. However, such is not the case. The data in Table 3 indicate this clearly. The data are from Caltrans classification counts collected routinely as part of an FHWA-mandated counting program and are for the six sites of the BioTech study at which the state made such counts. The doubles-to-singles ratio was greater at night than during the day in every instance except one--eastbound traffic at Site 122. Often the ratios differed by factors of three or more. By using only daytime data the exposure of doubles was greatly underestimated, relative to singles. Therefore, the doubles accident rate was greatly inflated relative to singles. Unfortunately, it is not possible to estimate how large an error was introduced overall because Caltrans counts are available for only six sites.

TABLE 3 Directional Differences in Truck Volumes

Site	Direction	Ratio of Doubles Volume to Singles Volume[a] Day	Night
112	N	0.354	0.975
	S	0.301	0.951
114	N	0.495	0.945
	S	0.599	0.825
115	N	0.343	0.458
	S	0.371	0.412
122	E	0.364	0.333
	W	0.063	0.250
141	N	0.236	0.920
	S	0.211	0.537
144	E	0.839	1.728
	W	0.937	0.990

[a]From Caltrans classification counts, July-September 1977.

Quality of Data

The problems with the exposure data discussed so far can be evaluated fairly objectively. They deal, essentially, with how the data base was created and what data are included. It is more difficult to assess the quality of the data. For the most part the truck classification counts are based on the judgments of technicians who viewed the photographs of vehicles. Because the photographs apparently have since been destroyed or discarded, it is not possible to confirm the accuracy of the judgments. Nevertheless, it is possible to examine the data base in great detail and note discrepancies that are highly suggestive of reading errors.

The film readers attempted to classify each vehicle photographed into one of eight categories:

- Straight truck,
- Straight truck plus full trailer,
- Straight truck plus dolly,
- Single,
- Double,
- Triple,
- Bobtail (tractor only), and
- Nontruck.

Figures 3 and 4 suggest how difficult the classification process can be. These are photographs taken under essentially optimum conditions by MRI at a number of the sites in California.

The difficulty of distinguishing vehicle classifications from photographs, particularly when they are poorly exposed (as at night) was recognized by BioTech (BioTechnology, Inc., unpublished progress reports to FHWA, December 1975). To deal with this difficulty, they established two additional vehicle classifications, "unknown large truck" and "unknown vehicle." The first was to be used when the film reader could not make a clear distinction between truck classifications; the second was to be used when, obviously, a vehicle had triggered the camera to take a photograph but, for whatever reason, it was not possible to tell if the vehicle was a car or a large truck. This might happen, for example, if there were no illumination or if another vehicle or object was between the lens and the target vehicle.

These difficulties were sometimes pervasive. Table 4, for example, indicates that at a number of

FIGURE 3 Similar-appearing tank trucks.

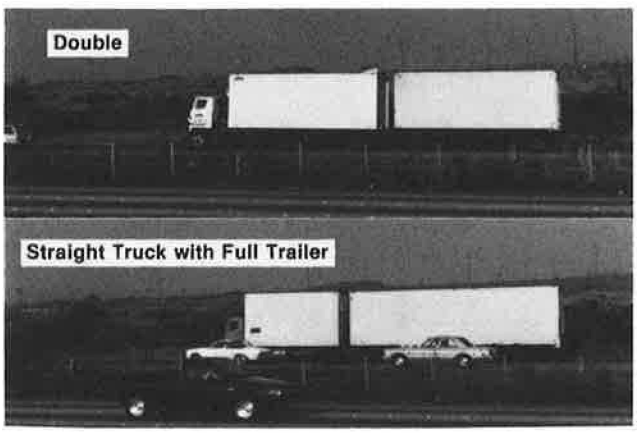

FIGURE 4 Similar-appearing van-type trucks.

TABLE 4 Sites with a High Percentage of Unknown Large Trucks

	Percentage of All Large Trucks			
Site	Singles	Straight plus Trailer	Doubles	Unknown Large Truck
111	33.3	12.5	11.2	26.3
113	58.0	14.7	15.7	6.9
114	52.2	17.2	19.0	7.2
115	40.9	17.1	12.8	22.4
121	29.1	25.2	26.8	10.1
131	29.8	19.0	4.8	2.6
133	30.1	6.5	34.2	12.6
142	29.8	21.7	14.9	8.1
144	33.1	21.4	25.0	13.8
152	40.4	20.0	13.2	15.9
162	3.9	7.9	3.9	3.9
413	65.4	8.6	15.3	4.0
421	54.1	4.4	5.1	26.1
442	44.0	12.3	6.2	10.9

sites the category "unknown large truck" contained the same magnitude of counts as the common truck classifications, and sometimes many more. For example, at Site 111 there were more unknown large trucks than there were doubles and straight trucks with trailers combined. After a review of much of the detailed count data, it appears that these unknowns were most likely either doubles or straight trucks with trailers. The net effect of this difficulty, then, is an underestimate of the exposure of doubles and, therefore, an overestimate of their accident rate.

Many specific, "unusual" data values were discovered on the data tape that were further evidence that data problems were encountered. A few brief examples follow:

1. Site 123, fifth quarter, night

Day	Nontrucks
Friday	23
Saturday	30
Sunday	57
Monday	3
Tuesday	28

2. Site 131, fourth quarter, day

Day	Straight Trucks
Tuesday	18
Wednesday	19
Thursday	3
Friday	4
Saturday	42
Sunday	27

3. Site 111, fourth quarter, night

Day	Unknown Large Trucks
Friday	5
Saturday	98
Sunday	0
Monday	1
Tuesday	4

4. Site 122, fourth quarter, day

Day	Straight with Trailer	Singles
Monday	1	43
Tuesday	5	20
Wednesday	34	0
Thursday	0	18

5. Various sites, day versus night

	Site 144		Site 421		Site 152	
Truck Type	Day	Night	Day	Night	Day	Night
Singles	137	24	602	50	151	8
Doubles	65	0	63	0	73	0
Straight	42	0	27	0	30	0
Unknown large trucks	29	87	4	253	6	36

The data in Table 5 suggest in a summary way the combined effects of the several difficulties inherent in the BioTech exposure data. The table covers all the sites for which comparable classification counts were available from Caltrans. Most of the latter were obtained in 1976 and 1977 and represent a complete count (100 percent sample) for a 24-hr weekday. (A few were obtained in 1975.) All Caltrans data are for both directions of travel except for Sites 131 (eastbound only) and 152 (southbound only). In every instance the BioTech data underestimate doubles exposure relative to singles exposure, compared with the Caltrans data. On the average, BioTech's data appear to underestimate doubles by 36 percent.

Size and Weight Data

The size and weight data base does not figure directly in the doubles versus singles issue. It does,

TABLE 5 Overall Underestimation of Doubles in Exposure Data Base

Site	Doubles-to-Singles Ratio		Caltrans-to-BioTech Ratio
	BioTech	Caltrans	
111	0.336	0.428	1.27
112	0.316	0.476	1.51
114	0.364	0.694	1.91
115	0.316	0.391	1.24
122	0.054	0.133	2.46
123	0.487	0.840	1.72
131	0.161	0.176	1.09
141	0.250	0.298	1.19
144	0.756	0.861	1.14
152	0.327	0.438	1.34

however, affect peripheral conclusions asserted by BioTech, such as that empty or nearly empty vehicles (especially doubles) have substantially higher accident involvement rates than do loaded vehicles. It is therefore enlightening to briefly review the size and weight data base.

Data Collection Sites

The intent was to collect size and weight data at the same sites that were used for accident and classification data collection. However, this was accomplished at only eight of the 31 sites. At these eight locations, all in California, size and weight data were obtained at state scales.

Data were also obtained at state scales at five other California locations, but not necessarily on the same highway as a site. No scales were available in Nevada, but one was used in Utah near the state border. Additional data were obtained from interviews at truck stops. Altogether the data base contains data on more than 27,000 trucks from California and Nevada obtained at a total of 28 locations. That the locations did not always match the study sites (that is, they did not sample the same traffic stream) is perhaps moot, however, because no site-specific analyses were performed. All analyses performed, and conclusions drawn, were based on the entire, pooled, data set.

Data were generally obtained for only one direction of travel, with the implicit assumption that the two directions were equivalent. At many California sites, however, that is simply not true. Weight data obtained by Caltrans in both directions at six of the BioTech sites illustrated this. The Caltrans data were obtained routinely as part of an FHWA-mandated truck size and weight monitoring program. Table 6 gives the median weight of doubles for the summer of 1977 (Site 123 data were obtained in 1975). There are logical reasons for these differences. For example, most of the doubles at Site 123 were flatbed trucks that carried agricultural produce from the fields to the canneries and returned empty. Most of the doubles at Sites 114 and 141 were tankers hauling petroleum products away from refineries and returning empty.

TABLE 6 Direction of Travel Versus Truck Weight for Doubles

Site	Directions of Travel	Median Weight (thousand lb)
112	North/South	38/59
114	North/South	54/76
115	North/South	64/66
123	East/West	28/74
141	North/South	34/78
144	East/West	50/76

Empty Bypass Effect

The truck scales in California employ a "bypass" lane. Trucks that are empty are directed, by sign, to use this lane and bypass the scales. Because these trucks were not weighed, the BioTech data base greatly underestimates the low end of the truck weight distribution. This was recognized by the authors, and a second set of calculations was performed assuming 25 percent of all trucks (doubles and singles alike) were empty.

The Caltrans weight data were obtained in surveys in which the bypass lane was closed. The most common truck configurations (e.g., 3S2) were sampled on a 25 percent basis and all others on a 100 percent basis. Figures 5-7 show the comparative weight distributions based on data from the eight California sites where BioTech obtained scale weight data.

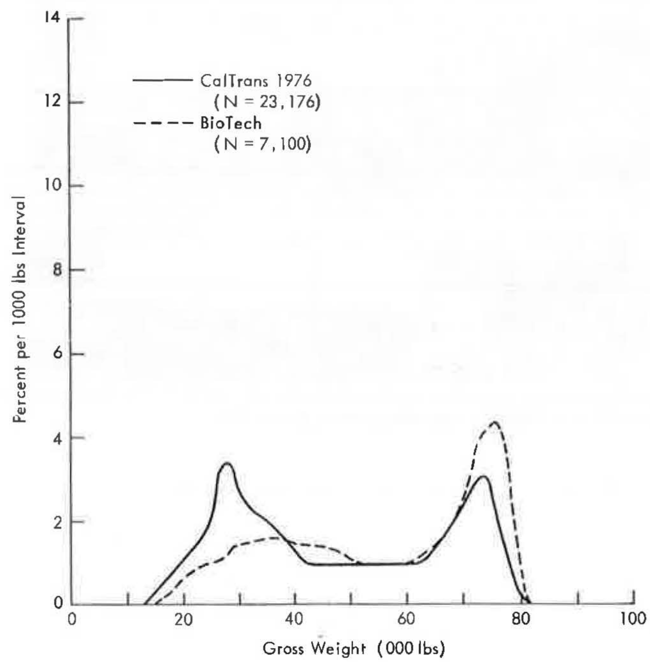

FIGURE 5 Weight distribution of singles.

Clearly, the discrepancies are sizable. The BioTech undercounting of empties results in an apparent overrepresentation of heavy trucks. Also, the discrepancies are not the same for the different truck classifications. Using 39,000 lb as an indicator of empty or not empty, the data in Table 7 are obtained. The data show that to correct the BioTech data base for undercounting empty trucks would require increasing their counts for empty singles by 68 percent, doubles by 395 percent, and straight plus trailer trucks by 614 percent. (Comparable results are obtained with other "indicator" values such as 33,000 lb.)

Missing Data

Of the entire California and Nevada size and weight data base, 93 percent of the data were obtained in

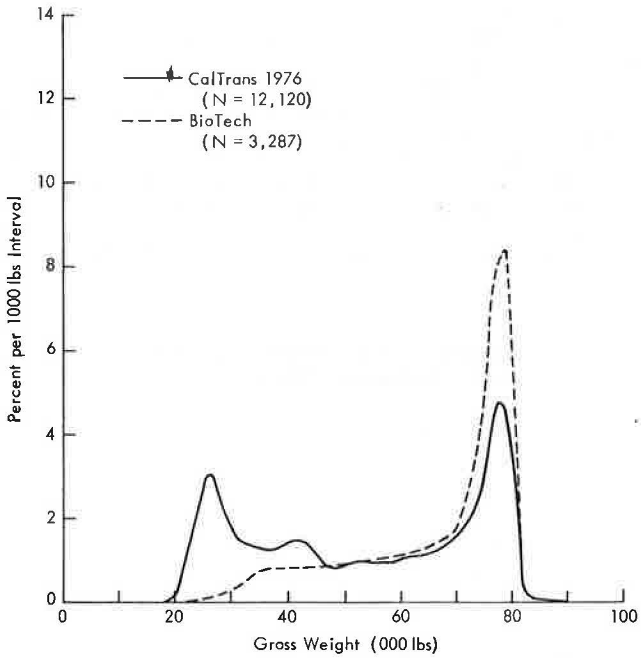

FIGURE 6 Weight distribution of doubles.

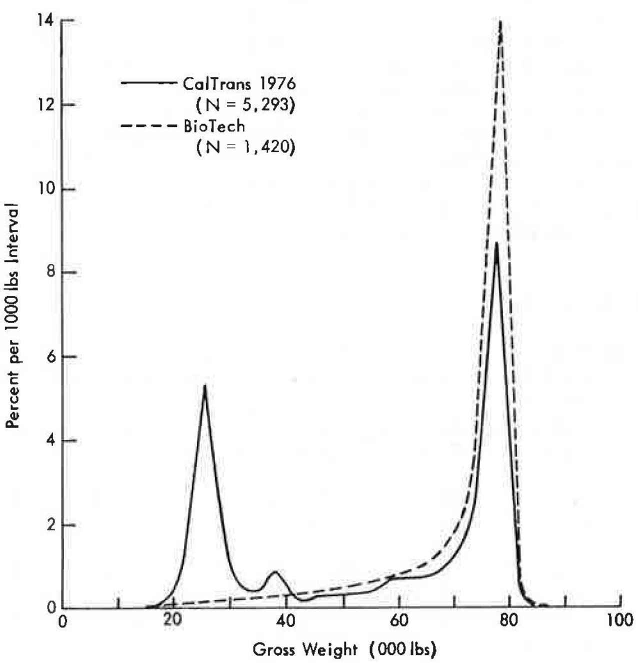

FIGURE 7 Weight distribution of straight trucks with trailers.

California. Nearly all (89 percent) were obtained during the day; it is not clear what biases this might have introduced. Presumably, for doubles, agricultural products would be less likely to be moved at night, and vans would be more likely then. Overall, according to Caltrans data, day and night truck volumes were comparable. However, doubles were relatively more dominant at night, as the data in Table 3 indicate.

In addition to truck weight, other data items were collected and analyzed, and conclusions were drawn. For example, one analysis compared accident rates of singles with 40- and 45-ft trailers. But this data element (length of first cargo unit) is missing for 86 percent of the trucks. A number of analyses examined driver parameters such as age and experience. These data are missing for 92 percent of the trucks.

TABLE 7 Undercounting of Empty Trucks

Classification	Percentage of Trucks Weighing 39,000 lb or Less: BioTech	Caltrans
Singles	26.3	44.3
Doubles	7.3	35.4
Straight plus trailer	4.4	31.4

Accident Data

It was intended that all truck accident involvements occurring on the study sites in the year-and-a-half study period be investigated in depth. A total of 2,112 accident involvements are in the data base. A review of those from California, in comparison with the Caltrans computerized accident file, revealed a fairly good match, considering the difficulty of collecting such data. More than 90 percent of the truck accidents in the Caltrans files were covered by the BioTech data base. An independent review of the accident data in the file from another of the six states in the study, Michigan, and the state police data files, indicated about 45 percent of the accidents were missing (J. O'Day, University of Michigan, unpublished data). However, the Michigan data were not used in the analyses of doubles versus singles, so any problems with the Michigan data do not affect that aspect of the study.

The accident data base consists of data taken from the police accident reports plus data obtained by follow-up investigations by the researchers. The files are reasonably complete (~ 95 percent) with respect to the police report data, but much of the expected follow-up information is missing. Missing data rates for these items range from about 30 percent (cargo area configuration), to 59 percent (truck weight), to 90 percent (truck weight distribution). The reasons for the missing data, such as difficulty in getting police reports in a timely fashion; difficulty in tracking down the involved drivers, trucks, or companies; and constraints of time and funds, are well understood by those versed in accident research.

The missing data of most importance, probably, are the truck weight data. These data were analyzed in comparison with other variables to determine if biases may have been introduced. Table 8 gives the differences in missing data rates as a function of cargo type. The cargo type "empty" had a known weight associated with it in more than 90 percent of the cases--that is, fewer than 10 percent were missing. Other cargo types had much more missing weight

TABLE 8 Likelihood of Knowing the Weight of an Accident-Involved Truck Versus Its Type of Cargo

Cargo Type	Percentage with Known Weight	N
Lumber products	50.0	18
Farm products	57.1	49
Solids in bulk	58.3	24
General freight	81.4	59
Liquids in bulk	83.9	31
Empty	90.2	143

information. More than 40 percent of the weights were unknown for trucks carrying farm products, a prevalent type of cargo in California. The effect of missing weight data for all truck classifications is evident in Table 9. The weight was more likely to be known if the truck was empty than if it was not empty (partly or completely loaded).

TABLE 9 Likelihood of Knowing the Weight of an Accident-Involved Truck Versus Its Empty/Not Empty Status

	Percentage with Known Weight	
Truck Type	Empty	Not Empty
All trucks	90.0	70.1
Single	95.6	81.5
Doubles	91.2	62.5
Straight	90.5	69.3
Straight plus trailer	84.6	69.6

The net effect of using the weight data with so much of it missing was to overstate the accident involvement rate of lighter trucks relative to heavy trucks, simply because the lighter (empty) ones were more likely to have a known weight than were the heavier (loaded) ones.

Data Analysis

The analyses presented in the BioTech report (10) were reviewed and many problems were found. The major problem, which affected all the computations, was the failure to account for missing data.

For example, the weights of about half the accident-involved trucks were not known. The computations of A_{ij} used only the trucks with known weights. Thus the calculated accident involvement role is apparently low by a factor of 2 because all of the vehicle-miles of travel were apportioned among only half the accident-involved trucks. In actuality, the situation is more complicated because the classification data (and the size and weight data) are also incomplete, and no corrections were made. Thus, for example, with 92 percent of the drivers' ages missing from the size and weight file (and assuming complete data on this item in the accident file), the computed rates would be high by a factor of 12.5. In summary, none of the computed accident involvement rates are numerically correct.

As indicated, a number of other problems were found in the analyses but need not be dwelled on. They could be corrected; the data bases cannot.

Representativeness of Truck Configurations

Even if there were no other problems with the BioTech study (10), a question would still need to be asked: "Are the results obtained likely to be observed elsewhere; that is, are they representative?" The authors clearly warn the reader not to extrapolate the data beyond the states in the study. However, many readers will be tempted to do so, and many have already (11-13).

The doubles versus singles issue, in most people's opinion, deals with the common van-type semitrailers and the increasingly common "twins," which are tractors plus two 27-ft van trailers. The latter, in particular, are becoming evermore popular in the "less-than-truck-load" (LTL) trucking industry--carriers involved in general commodity freight. Are these the types of trucks examined in the BioTech study?

Table 10 gives, for accident-involved doubles, the types of cargo configurations from the BioTech study. Of these 196 trucks, fewer than one-third (62) are of the common van or "twin" variety. There were more platform (flatbed) doubles than vans; such flatbed combinations are typically used to haul agricultural produce in California. Double tankers and bulk commodity carriers are also common in California.

TABLE 10 Cargo Area Configuration of Accident-Involved Doubles

Configuration	Number
Fully enclosed	62
Platform	74
Tank	30
Bulk commodity or dump	24
Other	6
Total	196

These nonvan doubles are frequently found in intrastate use in California but are not expected to become widely used elsewhere in the United States. Because of a quirk in California size and weight legislation before 1973, a double configuration could legally carry about 3,000 lb more than a single. This resulted from a kingpin-to-rear axle limitation coupled with bridge formula axle load limits. Thus, despite their greater capital investment requirements, there was an economic incentive to employ doubles in many industries in California. Although that particular economic incentive no longer exists, most of these industries continue to use such doubles within the state.

CONCLUSIONS

The MRI review of the BioTech study found that the major conclusions of that study are not supported by the project data base. The conclusion that doubles have substantially higher accident involvement rates than singles cannot be supported because

- Only one direction of traffic was sampled for classification data, but accidents in both directions were used. This is important because Caltrans data show that truck accidents and exposure differ greatly by direction at many sites.
- Most of the classification data were obtained during the daytime, even though about half of the accidents occurred at night. This is important because Caltrans data show that truck exposure, by type and configuration, differs greatly between day and night.
- Classification data were typically collected during only two quarters, not six as reported. This is important because both exposure and accidents differ greatly from season to season.
- The photographic classification data were based on a sample of axles, not vehicles, and no correction was made for this.
- The photographic classification data collected by BioTech differ greatly from manual classification data collected by Caltrans, even when comparisons are made between data collected at the same site in the same direction during the same quarter. These differences apparently arise from difficulties in interpreting the photographic data, which led to

a large undercounting of doubles and thus an inflated doubles accident rate.

The conclusion that empty or nearly empty trucks have high accident rates relative to loaded trucks cannot be supported because

- The truck weight data collected by BioTech differ greatly from those collected by Caltrans at the same scales, mainly because California routinely allows empty trucks to bypass the scales and these bypassed trucks were not sampled by BioTech. This problem resulted in a large undercounting of empty trucks, especially doubles and straight plus trailers, which led to inflated accident rates for empty trucks. No adequate correction for the bypassed trucks has been made.
- Most of the truck weight data were obtained during the daytime, even though one-half of the accidents occurred at night.
- Truck weights were missing from the accident data more often for loaded trucks than for empty trucks. This led to an inflated empty accident rate.

Many of the other conclusions of the study are not supported because of similar problems in the exposure and accident data bases.

Overall, the accident data base may be useful to future researchers, if care is taken to handle missing data properly. The exposure data bases, including both the classification data and the size and weight data, are totally inaccurate and analysis results derived from these data bases are probably meaningless.

ACKNOWLEDGMENTS

The research reported herein was funded largely by Consolidated Freightways, through the law firm of DeWitt, Sundby, Huggett and Schumacher, S.C. The authors are particularly indebted to John H. Lederer and William Lewis of that firm for their assistance and shrewd insight, as well as a thorough knowledge of trucking operations. The authors also acknowledge the help of Karin Bauer with some of the statistical analyses, Rosemary Moran and Debra Hodge for the great efforts required in the assembly and analysis of the data, and A.D. St. John for his keen insight on several of the matters reviewed.

REFERENCES

1. Surface Transportation Assistance Act of 1982, Pub. L. No. 97-424, 96 Stat. 2097 (1982).
2. J. Glennon. Matched Pair Analysis. Consolidated Freightways, Palo Alto, Calif., 1979.
3. R. Zeiszler. Accident Experience of Double Bottom Trucks in California. California Highway Patrol, Sacramento, 1973.
4. R.E. Scott and J. O'Day. Statistical Analysis of Truck Accident Involvements. Final Report. Highway Safety Research Institute, University of Michigan, Ann Arbor, 1971.
5. C.S. Yoo, M.L. Reiss, and H.W. McGee. Comparison of California Accident Rates for Single and Double Tractor-Trailer Combination Trucks. FHWA Report FHWA-RD-78-94. FHWA, U.S. Department of Transportation, 1978, 70 pp.
6. K.L. Campbell and O. Carsten. Fleet Accident Evaluation of FMVSS 121. NHTSA, U.S. Department of Transportation, 1981.
7. T. Chirachavala and J. O'Day. A Comparison of Accident Characteristics and Rates for Combination Vehicles With One or Two Trailers. Report UM-HSRI-81-41. University of Michigan, Ann Arbor, 1981, 92 pp.
8. T. Chirachavala and J. O'Day. A Safety Comparison of Doubles Versus Tractor-Semitrailer Operations. FHWA Technical Report. Bureau of Motor Carrier Safety, U.S. Department of Transportation, 1977.
9. L.L. Phillipson et al. Statistical Analyses of Commercial Vehicle Accident Factors. NHTSA Report DOT-HS-803-418. NHTSA, U.S. Department of Transportation, 1978, 400 pp.
10. G.R. Vallette et al. The Effect of Truck Size and Weight on Accident Experience and Traffic Operations, Vol. III: Accident Experience of Large Trucks. FHWA/RD-80/137. BioTechnology, Inc., Falls Church, Va.; FHWA, U.S. Department of Transportation, 1981, 145 pp.
11. G.R. Vallette et al. Report to Congress on Large Truck Accident Causation. NHTSA, U.S. Department of Transportation, 1982.
12. M. Freitas. Review of Accident Research Involving Truck Size and Weight. Public Roads, Vol. 46, No. 2, 1982.
13. M. Freitas. Influence of Truck Size and Weight on Highway Crashes. Status Report, Vol. 18, No. 4, Insurance Institute for Highway Safety, Washington, D.C., March 8, 1983.
14. M. Freitas. Technical Evaluation of the BioTechnology, Inc., Study, The Effect of Truck Size and Weight on Accident Expansion and Traffic Operations. Transportation Systems Center, U.S. Department of Transportation, 1981, 60 pp.

Publication of this paper sponsored by Committee on Motor Vehicle Size and Weight.