5.0 Procedures and Methods

5.1 Safety instructions to drivers regarding fatigue management technologies

Drivers (and company officials, since they own the trucks) were told that the *Sleep Watch*®, the *Copilot*® (automated PERCLOS), and *SafeTRAC*® were prototype systems that offered promise as monitors of driver fatigue (the fourth component of the FMT combined system was the *Howard Power Center Steering® (HPCS)* system, which is not a prototype, but a marketed device that is already in use in some trucking and motorcoach operations). Drivers were instructed that the investigators were interested in their experiences and opinions of these systems (i.e., during the 2-week period of the FMT FEEDBACK condition when theses devices were providing feedback to drivers and when the *HPCS*® was engaged). Since three of the devices were prototypes, drivers were informed that the digital displayed feedback indications (e.g., lighted display numbers from 0 to 99 provided by the *Sleep Watch*®, the *Copilot*®, and *SafeTRAC*® technologies) may not match their sense of how tired or alert they really are, and that they should use their own professional judgment regarding their alertness, fitness to drive, and need for rest, always staying within the applicable Federal hours-of-service.

Drivers were instructed to use the fatigue management technologies in a responsible and safe manner. The Informed Consent expressly stated, “Drivers are responsible at all times for managing their own levels of fatigue and alertness. Drivers must assess their own condition and use their own judgment rather than rely on the devices to make decisions about whether or not to drive.” It was emphasized in both conditions (NO FEEDBACK and FEEDBACK) that if they felt they were unfit to drive for any reason, they should stop driving, regardless of what the technologies indicated or the schedule or regulations permitted. Drivers therefore remained the ultimate arbiters of their ability to drive safely. The research team worked closely with companies and drivers to ensure that everyone involved understood that a driver should terminate driving if he or she felt unable or unfit to drive for any reason, in any condition (i.e., FEEDBACK or NO FEEDBACK). Thus, this study did not involve any explicit (or implicit) encouragement of drivers to violate the hours-of-service in country in which the drivers worked.

This pilot study also did not manipulate or control work-rest schedules of drivers. Thus, no effort was made to control the development of driver "sleep debt," which was not the focus of the study—any sleep debt experienced by drivers would be whatever normally occurred within their chosen lifestyle, within the operations of their company employer, and within what is possible through application of the hours-of-service in Canada or the U.S. Instead, the focus of this pilot study was on the extent to which FMT FEEDBACK altered lane tracking, slow eyelid closures, and sleep obtained within what Canadian hours-of-service or U.S. hours-of-service would allow.

In addition to volunteer drivers participating for 4 working weeks of data collection (2 weeks with NO FEEDBACK followed by 2 weeks with FMT FEEDBACK), they also received at the beginning of the study 3 hours training on "Alertness and Fatigue Management" as well as training on each FMT hardware system deployed. They also participated in two structured human factors debriefing interviews (one at the end of the 2-week NO FEEDBACK period and a final one after the 2-week FMT FEEDBACK period).

5.1.1 Fatigue education module

Education on Alertness and Fatigue Management was provided to all drivers enrolled in the study, after informed consent was given but before they drove the FMT-instrumented trucks. Drivers were provided an approximately 3-hour course entitled "Mastering Alertness and Managing Driver Fatigue," (sponsored by Federal Motor Carrier Safety Administration and the American Transportation Research Institute), which was prepared for this study and taught by Dr. G. Krueger of Krueger Ergonomics Consultants (see Appendix B-1). The 3-hour course was taught to four drivers at a time, 2-3 days before they were issued their instrumented trucks and before beginning their 4-week participation in the data collection portion of the study. The Education Module encouraged drivers to be responsible for their alertness levels and to use the information they gained through the module to their benefit.

5.1.2 Confidentiality of data

As mandated by the Human Subjects Ethics Committees in both Canada and the U.S. that reviewed and approved the protocol, participant drivers were informed that their data records from the study were kept confidential from their company and parties other than the project investigators. Each driver's identity was codified with a unique ID number. The investigators did not report driver behavior to the company, but company dispatchers and driver managers were aware of participant drivers' work-rest behaviors, as they are routinely.

5.2 Data acquisition procedures

In the Canadian phase of the study, four Challenger Motor Freight trucks—all with single trailers—were instrumented with the FMT equipment and rigged for data collection. This included two each of the following kinds of trucks: Volvo Base Model VNL64T with HSS660 Full Integral Sleeper Cab, and Freightliner Model CST420 Conventional Chassis, set back front axle tractor, also with a sleeper berth. Challenger Motor Freight is an ISO 900 company and also has been rated as one of Canada's 50 best managed companies. It offers a wide array of transportation services, from truckload shipments across the continent and into Mexico, to local same day service, to warehousing and inventory management. All trucking operations were out of Challenger's terminal at Cambridge, Ontario, Canada.

In the U.S. phase of the study, four Con-Way Central Express trucks were instrumented with FMT equipment and recording equipment for data collection. They were all Sterling Model MN-80 trucks with double trailers. Con-Way Central Express (CCX) is a regional motor carrier providing next and second-day service throughout 25 Midwestern and Northeastern states of the U.S. CCX is one of seven businesses that are part of Con-Way Transportation Services, Inc. Con-Way is a $2 billion company that provides time-definite and day-definite freight delivery, and logistics services for commercial and industrial businesses. See Appendix A-1 for photos of trucks used in both Canadian and U.S. study phases. Conway's participating drivers departed and returned nightly from terminals at Hermitage, Pennsylvania or Lordstown, Ohio.

As summarized above, the FMT system instrumented on all U.S. and Canadian trucks involved four technologies: (1) *SleepWatch®* with Sleep Management Model; (2) The *Copilot®* (automated PERCLOS monitor); (3) *SafeTRAC®* lane tracker; and (4) *Howard Power Center Steering®* (*HPCS)* system. The first three technologies also recorded data without providing feedback to drivers. Therefore, they were used in both of the 2 weeks of NO FEEDBACK (control) condition, and the 2 weeks of FMT FEEDBACK (intervention) condition. Data used in hypothesis testing was obtained from the following four devices.

5.2.1 *Accident Prevention Plus* (*AP+*) on-board recording device

Trucks were instrumented with the *AP+ black box* on-board recording device to provide data every second the truck was running throughout the approximately 1 month period each driver was studied (see footnote 5). Figure 6 displays the *AP+ black box* recording unit (see also Appendix A-1). The *AP+ black box* continuously recorded (every second for the 4 weeks of the study) measures derived from the following variable domains.

- Drivers' levels of alertness-drowsiness while driving based on measurements made by the *Copilot®,* which consisted of infrared detection of slow eyelid closures and a proprietary algorithm that yielded a numerical value for PERCLOS (Appendix B-1).

- *Copilot®* feedback information on drowsiness as presented to the driver (only in the FMT FEEDBACK condition). These were recordings of the digital readout displayed to the driver representing his/her level of drowsiness from PERCLOS (percent slow eyelid closure)—Figure 3.

- Lane tracking performance variables from *SafeTRAC®* (e.g., lane tracking score, and lane displacement), and drivers' levels of alertness while driving based on measurements made by *SafeTRAC®* and integrated in its proprietary algorithm.

- *SafeTRAC®* feedback information on driver alertness (only in the FMT FEEDBACK condition).

- *AP+* recorded steering performance variables from strain gauge sensors mounted on the steering column and on the axle-wheel combination. This was done to evaluate the impact of the *Howard Power Center Steering®* on driver steering performance.

- ***AP+*** was attached to trucks in a manner that permitted recording of driving performance variables that included speed; braking; and lateral acceleration.

- ***AP+*** also recorded time-of-day and ambient outdoor light.

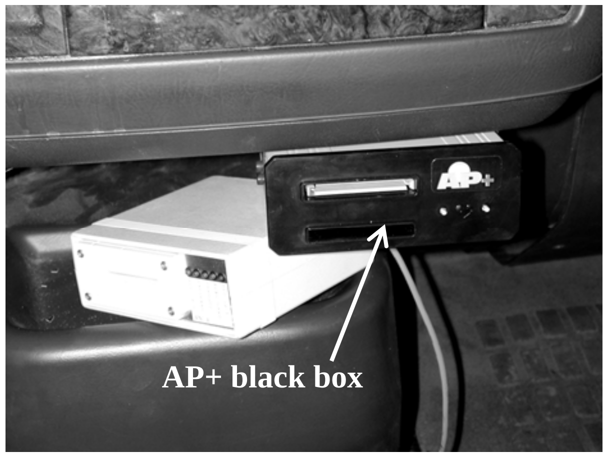

Figure 6. *AP+® black box* recording device mounted under truck dash. (See Appendix B-1 on instructions to drivers regarding use of the AP+ black box recording system.)

5.2.2 Portable *Psychomotor Vigilance Task* (*PVT-192*) reaction time device for measuring behavioral alertness

The *PVT-192* device was not considered by investigators as an FMT technology, but rather it was provided to drivers for measurement of behavioral alertness levels. Thus, the PVT was used as an independent objective evaluation on drivers' alertness-sleepiness. It does this by assessing drivers' reaction times during a 10-minute visual vigilance task. Drivers were asked to complete the PVT 10-minute reaction time vigilance test at the midpoint and end of each driving workday. The PVT is a well-validated 10-minute laboratory test of behavioral alertness that is widely used to obtain an estimate of performance limits in alert and drowsy subjects.[9] It was developed and extensively validated through scientific research by D.F. Dinges and colleagues.[10] The *PVT-192* is the portable version of the task (Ambulatory Monitoring, Inc. Ardsley, NY). Appendix A-1 shows two sketches of the portable PVT-192 unit used in the study, and Appendix B-1 contains the instructions given to drivers regarding self-administration of the PVT at the midpoint and end of each driving workday (drivers only did the 10-minute test while the truck was stopped).

9 Dorrian, J., Rogers, N.L., Dinges, D.F.: Psychomotor vigilance performance: A neurocognitive assay sensitive to sleep loss. In Kushida, C. (Ed.), Sleep Deprivation. Marcel Dekker, Inc., New York, NY.

10 Dinges, D.F., Powell, J.W.: Microcomputer analyses of performance on a portable, simple visual RT task during sustained operations. Behavior Research Methods, Instruments, & Computers 17 (6):652-655, 1985. Jewett, M.E., Dijk, D.J., Kronauer, R.E., Dinges, D.F.: Dose-response relationship between sleep duration and human psychomotor vigilance and subjective alertness. *Sleep* 22 (2):171-179, 1999.

5.2.3 *SleepWatch®* actigraphic technology to record drivers' sleep/rest times

Information on drivers' "performance fuel gauge," which was the product of an algorithm based on sleep/rest times, and information on drivers' "performance readiness" level (P = from 0% to 99%) were acquired from the Walter Reed Army Institute of Research *SleepWatch®* (Figure 1; Appendix A-1 and Appendix B-1). The *SleepWatch®* also provided basis actigraphic data every minute across the 4-week period of the study for each driver, which was used to estimate the sleep obtained by drivers during the NO FEEDBACK and FEEDBACK conditions. Such information served as a primary outcome for hypothesis testing.

5.2.4 *Daily diary booklet* for drivers to record driving conditions, work, naps, etc.

Drivers were provided a daily diary to record driving conditions (weather, slow traffic; hilly roads, crosswinds, waiting); work activities (loading and unloading; deliveries; etc.); rest breaks and naps; days off; reactions to FMT devices; and day and night activities (work, rest, sleep). These diaries were used to enrich understanding and interpretation of the objectively recorded data from trucks (via *AP+®*) and drivers (e.g., *SleepWatch®* and *PVT-192*).

5.2.5 Post-study *Human Factors Structured Interview Questionnaire*

Drivers were administered a *Human Factors Structured Interview Questionnaire* (developed by G. Krueger) at the end of the 2-week NO FEEDBACK period, and again at the end of the 2-week FEEDBACK condition period. The questionnaire was administered in a structured interview debriefing session. It asked drivers to answer specific questions and provide their perspectives on the following interventions: *Alertness and Fatigue Management Training Course*; *SleepWatch®*; *SafeTRAC®*; *Copilot®*; *Howard Power Center Steering®*; *Psychomotor Vigilance Task (PVT)*; and the combined *Fatigue Management Technologies* used in the study.

5.3 Defining data records for statistical analyses

Given the extraordinarily large volume of data gathered in the study, it was necessary to determine data management and variable extraction procedures that would ensure quality control of the data. Of particular concern was the need to utilize procedures that avoided including erroneous data values (especially data corrupted by equipment failure in the field—it is important to keep in mind that while all the equipment accompanied drivers during 4 weeks of work, no investigator or study technicians were present while drivers were on the road, and hence no one was present to prevent data loss or corruption from equipment damage due to the environmental conditions [e.g., vibration, heat, cold, rain, snow and ice] in which it was deployed).

The data gathered from the *AP+®* black box recorder onboard trucks in Study Phase 1 (Canada trucks) was used to illustrate the processes and procedures put in place to ensure data quality control. (Keep in mind that the AP+ data recorded every minute for 4-weeks of driving in every subject, was only a subset of the data—other large data sets were obtained from the *SleepWatch®*, *PVT*, Daily Diaries, and Human Factors Questionnaire.) The actual results for the Canadian study phase are presented later in section 6.0 of the report. In the following sections we detail data handling procedures and statistical methods that were applied to both Study Phase 1 (in Canada) and Study Phase 2 (in the U.S.).

In order to assess quality of the data arising from the two study phases, the same three hierarchical sample definitions were used to construct summary Figures. These three sample definitions were required to manage the analysis of the very large volume of data recorded by the *AP+®* system.

- All *AP+®* data with no records excluded (6,472,457 total records among 26 drivers in Study Phase 1 in Canada, and 2,265,248 total records among 12 drivers in Study Phase 2 in the U.S.).
- *AP+®* data records in which speed was at least 30 mph (5,060,743 total records among 25 drivers in Canada, and 2,013,942 total records among 12 drivers in the U.S.).
- *AP+®* data for speed ≥ 30 mph, artifacts eliminated and records within measurement range (4,748,278 total records among 20 drivers in Canada, and 1,935,577 total records among 9 drivers in the U.S.).

These hierarchical data sets from the Canada trucks are briefly described below to illustrate our data handling techniques.

5.3.1 *AP+®* data with no records excluded

The first sample of 6,472,457 total records from 26 Canadian drivers included all raw data and was constructed for comparison purposes only. A limited number of Figures for which the entire sample is relevant show data from all records.

5.3.2 *AP+®* data in which speed was ≥ 30 mph

The second sample definition eliminated records in which speed was recorded as less than 30 mph. This left 5,060,743 total records among 25 Canadian drivers. There were two reasons for this exclusion. The primary reason was that the study was designed to examine the effects of the fatigue management intervention in highway driving and ≥ 30 mph was the definition adopted by the study team for highway driving. The second reason was that most records with artifacts were eliminated by restricting attention to records recorded when vehicle speed was at least 30 mph. Appendix C-1 Data Quality Control (DQ) Figure 1 for the Canada study phase data provides the number and percentage of all *AP+®* records recorded with speeds of at least 30 mph, less than 30 mph but greater than 0.62 mph, and equal to 0.62 mph (Appendix C-2 has the comparable analyses for the U.S. study phase). The value 0.62 mph was derived from conversion of 1 km/h, the value record by the AP+ system when the truck was standing still.

5.3.3 *AP+®* data in which speed was ≥ 30 mph, artifacts removed and records within measurement range

Careful examination of driver specific distributions of outcome variables recorded by the *AP+®* system in Study Phase 1 identified additional artifacts and problematic data among the remaining records with speeds at least 30 mph. These records were excluded on a case-by-case basis after careful evaluation, leaving 4,748,278 total records among 20 Canadian drivers. Reasons for exclusions of these records are documented in Appendices C-1 for Canada Study Phase 1 (see Appendix C-2 for similar data cleaning for the U.S. Study Phase 2). In addition, data from 6 Canadian drivers were not included in the clean analysis sample because *AP+®* recorded data was only available under one of the two conditions (NO FEEDBACK or FEEDBACK), or because there was insufficient data under one of the two conditions to permit meaningful comparisons. ***Thus, final cleaned analysis samples from both Canada and the U.S. were defined on the basis of the subset of drivers with sufficient data under both conditions (FEEDBACK and NO FEEDBACK), restricting attention to records recorded at speeds of at least 30 mph, after excluding additional data found to be invalid, following careful examination of driver specific distributions*.**

5.3.4 PVT data

The drivers included in the cleaned analysis sample as defined above were utilized in analyses of the PVT performance data for consistency across data sources.

5.4 Statistical methods

5.4.1 *AP+®* recorded outcomes

For each outcome variable recorded by the *AP+®* system, four analyses were performed to assess if there was a significant change from the NO FEEDBACK condition to the FEEDBACK condition within the Study Phase 1 in Canada, and again within Study Phase 2 in the U.S. These statistical methods are described below.

5.4.1.1 Unweighted analysis for means and standard deviations

The first analysis was implemented by computing unweighted mean values and standard deviation values across all records for a specific driver under a specific condition (NO FEEDBACK and FEEDBACK). Mean values were compared for the following outcome variables:

- *Copilot®* measures of PERCLOS during night hours mean (via *AP+®*)
- *SafeTRAC®* Driver's Alertness mean (via *AP+®*)

Standard deviations were compared for the following outcome variables:

- *AP+®* Lateral distance standard deviation
- *AP+®* Steering wheel movements standard deviation
- *AP+®* Front wheel movements standard deviation

Then within-driver change scores were computed. Paired t-tests were performed to assess the statistical significance of the changes in means or standard deviations as appropriate.

5.4.1.2 Weighted analysis

The second analyses introduced two weighting factors. First, when computing the within driver and condition mean, median, standard deviation, and interquartile range values, records were replicated if they corresponded to more than 1 second in duration. In this way, for example, records with durations that were 3 seconds contributed a weight 3 times greater than records with durations of 1 second.

Even accounting for record duration, drivers varied greatly with regard to the total duration of data in the cleaned analysis sample. Drivers with greater total durations under both conditions contribute more information with regard to intervention effects. In contrast, a driver with a short duration under one of the conditions contributes less information about within driver changes. To account for this, and to optimize the ability to consider both within-subjects and between-subjects sources of variance, mixed model analyses of variance were used to compare mean (duration-weighted) values between the NO FEEDBACK and FEEDBACK conditions, weighting by the total number of available records (separately by condition). This is referred to in the Figures in the Results section as the "doubly-weighted" analyses. All mixed model analyses were implemented using the Proc. Mixed procedure available in SAS[11].

5.4.1.3 Unweighted analyses for median or interquartile ranges

The analyses were repeated that summarized the NO FEEDBACK and FEEDBACK distributions of *Copilot®* PERCLOS during night hours and *SafeTRAC®* Driver's Alertness by median values rather than mean values, in order to provide summaries of the center of these distributions that are less sensitive to outliers and skewness. Similarly, *AP+®* Lateral distance, *AP+®* steering wheel movements, and *AP+®* front wheel movements were summarized using interquartile ranges (IQR) instead of standard deviations. The IQR is defined as the difference between the 75th percentile value and the 25th percentile value) and is less influenced by extreme values than the standard deviation. Both the paired t-test and mixed model weighted analyses were performed on the median and the interquartile range for each variable (which are the nonparametric alternatives to the mean and standard deviation).

5.4.2 Psychomotor vigilance task (PVT) outcomes

Mixed model analyses of variance[12] was used to assess the significance of the intervention effect (NO FEEDBACK vs. FEEDBACK), controlling for time-of-day category (day, evening, night). The initial model included fixed effects for time-of-day (morning, evening, night), presence vs. absence of feedback, and time-of-day by feedback interaction. It also included a random effect for driver to account for correlations within driver.

5.4.2.1 Reliability assessment using adjusted ICC

The interaction model (i.e., feedback condition, time-of-day, time-of-day by feedback condition) was used to compute an adjusted intraclass correlation (ICC). The intraclass correlation is the proportion of total variance explained by systematic differences among drivers after accounting for time-of-day and feedback condition effects.[13] PVT performance outcomes (e.g., median RT) are known to have relatively high ICC values (i.e., stable inter-individual variance). Assessment of ICC was taken as a quality control procedure and to document

---

11 SAS Institute Inc., SAS OnlineDoc®, Version 8, Cary, NC: SAS Institute Inc., 1999.

12 Burton P, Gurrin L, Sly P. Extending the simple linear regression model to account for correlated responses: An introduction to generalized estimating equations and multilevel mixed modeling. *Stat Med* 1998;17:1261-91. Van Dongen HPA, Olofsen E, Dinges DF, Maislin G. Mixed-model regression analysis and dealing with inter-individual differences. In: Johnson ML, Brand L, eds. *Numerical Computer Methods* (part E). San Diego: Academic Press 2004:139-171.

13 Fleiss JK, *The Design and Analysis of Clinical Experiments,* New York: John Wiley & Sons, 1986, Chapter 1.

the ability of this study to obtain reliable PVT performance assessments in the field.

5.4.2.2 The test for time-of-day by feedback interaction

The same model used to determine the ICC's was used to examine whether differences between responses obtained during the NO FEEDBACK and FEEDBACK conditions varied by time-of-day. A p-value of 0.10 was employed because of the low power inherent in tests for interaction.

5.4.2.3 The simple and main effects

If $p \geq 0.10$ then the interaction terms were removed from the model and the feedback effects and time-of-day effects were tested as main effects in the ANOVA model. If $p < 0.10$, we concluded that differences between the NO FEEDBACK and FEEDBACK conditions significantly varied by time-of-day. Therefore, separate mixed models were used to test for feedback effects at each time-of-day interval (day, evening, night).

5.4.3 *SleepWatch®* (actigraphy + mathematical model) outcomes

Daily mean values were analyzed for variables derived from the *SleepWatch®*. Mixed model analyses of variance were used assess the significance of the fixed intervention effect. Random effects included between and within driver variance, which were used to compute intraclass correlations.

5.4.4 Daily sleep diary and Human Factors Structured Interview Questionnaire

Descriptive statistics were used for analyzing the drivers' daily diary and post-experimental responses to the Human Factors Structured Interview Questionnaire.